

What to make of the price of oil

The failure of the oil-producers to agree last week has triggered off a wave of anxiety among their customers. Nervousness is understandable. Failure to grasp the opportunities will be unforgivable.

THE Organization of Petroleum Exporting Countries (OPEC) is in disarray, but not yet dead. But the failure of the oil exporters to agree on prices and production quotas at last week's meeting in Geneva is less remarkable than the panicky reactions of their customers. Bankers with loans outstanding in oil-producing states such as Mexico and Nigeria are understandably nervous. But money seems to have been sloshing around needlessly in the currency markets, while governments in industrialized states have been behaving as if the now modest fall in the price of oil is almost as great a calamity (for them) as the sudden threefold increase of the price of oil in the second half of 1973 — the event that ushered in the recession that has dragged on ever since. So far, little has been heard from those who take the calmer view that if an increase of the price of oil ten years ago was a disaster for some, a decrease now is almost certain to be the opposite — a blessing, if only a modest one, or at least an opportunity.

Although OPEC is commonly described as an oil cartel, in reality it is nothing of the kind. It was formed in the late 1950s, when the price paid by the industrialized West for crude oil from the Middle East was outrageously too low. In those days, an oil-producing state was lucky to collect 50 cents on a barrel of oil and the total cost of crude to the multinational oil companies was often as little as \$1.50 a barrel — a third or less of the then going rate for crude oil lifted in the United States. For the first decade of its existence, OPEC made little headway towards higher prices. In retrospect (but even at the time) the multinational oil companies and the oil-consuming governments were shockingly slow to appreciate that Libya's nationalization of its oil in 1969 would provide a springboard for the dramatic events of 1973. In reality, once the major sources of Western oil imports had legitimately been nationalized, and for as long as OPEC was the predominant source of internationally traded oil, the market price was in effect arbitrary, and was so fixed. What has happened in the past few years is a simply that OPEC dominance has been undermined partly by increased production from other sources (Alaska, Mexico and the North Sea, for example) but primarily by reduced demand for OPEC oil. Since the last substantial price increase in 1978, OPEC production has fallen by a half to just over 16 million barrels a day, but production elsewhere, outside the Soviet Union and Eastern Europe, now exceeds 20 million barrels a day (in round numbers, one barrel a day is 50 tonnes a year). During this period, OPEC has been powerless to moderate the market-driven decline of prices for lack of the two essential weapons of a cartel — effective control of supply, and an agreement among its members that production will be held down.

Collapse

So will the price of oil now plummet back to where it was in 1973, or fall as quickly as it then rose? This seems to be the fear of some of those now rushing to convert oil-backed sterling into other currencies. The calculation is mistaken, for even if OPEC did not exist, the international price of oil would be determined by the marginal cost of production from oil fields elsewhere, which may be as great as \$20 a barrel in some places in the North Sea and (counting pipeline costs) Alaska. At least so long as the members of OPEC sensibly seek to sell what oil they can at the best prices they can get for it, there is no danger that these marginal sources will be undercut in price, and thus rendered useless investments.

The most likely outcome of the present upheaval is that the price of oil will continue to drift downwards, but by no more than a third of the official price of \$32 a barrel for which Saudi Arabia was pushing at Geneva a week ago. For the chance is small that OPEC members will be able to settle their genuine economic differences in time for the meeting arranged for later this month: producers such as Iran, Nigeria and Venezuela need the income that comes from selling oil, and would need regular subventions from the only low-cost producers left (Saudi Arabia and Kuwait) before they could agree among themselves on production quotas. There is always a chance that OPEC may be stiffened in some such way, but it is more likely that market forces have returned for good.

Resurgence

The consequences of that development could be far-reaching, at least if oil-producing governments respond imaginatively. The increase of the price of oil in 1973 was calamitous for the industrialized oil consumers because it provoked an unprecedented blend of inflation (prices of a widely used commodity were sharply increased) and deflation (resources were taken out of the economies of oil consumers and transferred to those of the oil producers). Over the past decade, the true deflationary effects of the increased price of oil have been further enhanced by the high cost of investing in other sources of oil and alternative sources of energy as such. For although such investments are eventually productive, and their product counts towards a country's gross national product, the resources must come from somewhere, usually from personal consumption. The potential advantage of the new circumstances that have arisen is that it should now be possible to escape from that cleft stick. In most places, it should now be easier to keep inflation down, while resources that have hitherto been paid over to the oil producers will be partly released for making good the past ravages of deflation. That is why the break in oil prices is an opportunity.

There are, of course, some snags. Not all countries will be similarly affected. But there can be no question that the economies that have been most seriously affected in the past decade, the developing countries without oil resources of their own, can only benefit from a reduction of the price of oil. That is something to be grateful for. Countries such as Mexico, encouraged by the events of the past decade to invest borrowed funds both in oil production and in the anticipation of the general economic and social development that would normally follow, will however be in even more serious trouble than in the past few months. Even though the International Monetary Fund has now produced \$4,000 million (plus the cooperation of the commercial banks) to help Mexico out of its recent pickle, there is every reason why industrial governments should be prepared to dig deeper into their pockets in the months ahead. Countries such as Venezuela, in hock to the West but a full member of OPEC, present more difficult problems, substantially political. There is no ground for fearing that they are insoluble.

The industrialized countries are in a different case. To the extent that they are bound together by trade, and likely to be held in check by the huge deficit in prospect in the United States, with the high interest rates that seem to follow, they may see in the months ahead the substitution of one problem (the price of oil) by



another (the deficit). But the United States economy is at least flexible enough for the domestic price of oil already to have begun to fall (and for the pace of investment in oil production to have fallen, thus releasing resources for other purposes). Britain will have to deal with more perplexing problems. North Sea oil, once thought to be an uncomplicated boon, has become a crutch for the government's revenues. The temptation will be for the British government to make up for the reduced taxation that lower petroleum prices would bring by increasing excise taxes on the sale of oil products, on the principle that if people have become used to paying so much for fuel for their motor cars, they will keep on doing so even when the true price falls. That would be a damaging mistake, for the only way of escaping from between the millstones of inflation and deflation is to wring the full benefit from the chance to let oil products cost less.

Just how big those benefits will be is anybody's guess. In most industrialized countries, something like 5 per cent of expenditure goes on petroleum products. On the face of things, even a reduction of the price of oil by a third, the most that can be expected, should make only a marginal difference. But everything depends on where the liberated resources would be spent. Countries such as Britain, beggared for years by too little investment in plant and machinery, could be wonderfully transformed if only a fraction of the reduced price of oil went in that direction. Comparable countries in Western Europe (not to mention Japan) are unlikely to have serious doubts on the subject. The British case, however, is complicated by North Sea oil, and by the decline of the value of sterling against the US dollar in the past few months. The immediate result is that the price of petrol has actually increased because even North Sea oil is priced in dollars. But that should be only a passing phenomenon: the price of British oil has not yet been reduced, sterling will probably right itself as panic passes and the British government will take less in oil taxation if it wishes British industry to share in whatever economic revival may follow. For that is the prize. Even if OPEC's failure last week is only temporary, so that the oil-producers return to Geneva at some stage as a fully-fledged cartel, it is not too soon to ask what lessons should be learned from the past decade. The first of these is the need for realism even among the most sophisticated of economic states. In the 1950s, the industrialized West built its rapid growth on the assumption that dollar-a-barrel oil was an inexhaustible resource. Then in 1958, the late Sir Harold Hartley produced a report for OECD which argued a case against nuclear power on the proposition that oil was cheap and plentiful. Physically, the resource has lasted well, but events have shown that a raw material which is drastically and inequitably underpriced can be abolished at the stroke of a pen: OPEC's success in 1973 was not to abolish oil, but merely cheap oil. If prosperity returns, there will no doubt again be moves by the mineral-exporting poor countries to fix the price of materials such as copper at levels more like those obtaining on the international markets.

Oil-dependence as such has also proved to be dangerous. In the early 1970s, Japan and Denmark were, for example, almost totally dependent on oil for energy. They have learned painfully the need for diversity in the supply of strategic raw materials (not oil itself, but the joules that can be had from it). For most of the past two decades, one of the strongest arguments for the development of the technology of nuclear power has been the opposite of Sir Harold Hartley's: that the only strategic defence against arbitrary price-fixing for essential raw materials is to show that even they can be priced out of the market. What this implies is that budgets for energy research and development (as well as conservation) should not be slashed now that the price of oil is abating.

Finally, governments must by now have learned not to avoid economic reality for political reasons. Throughout the past decade, the price of natural gas in the United States has been kept artificially low. The lesson of the OPEC affair is that in the long run the only sensible price for some such commodity is the market price. So will the administration bite the bullet, and let gas prices find their own level? There may never be a better time.

Reagan's conversion

The US Administration has been won over by science and technology. What will Congress say?

PRESIDENT Ronald Reagan seems well along the road to a kind of technological conversion. His State of the Union message to the new Congress last week included several references to the importance of making a technologically strong society even stronger in this respect, while his plea that mathematics and science should be taught well to all students younger than sixteen was clouded only by his advocacy of the constitutional amendment to permit religious teaching in schools, certain to distract any member of Congress interested in education from mundane matters such as earmarked block grants from the federal government. The President's third budget (see opposite) on the face of things seems to reinforce this commitment to technological affairs — spending on research and development would increase by nearly \$7,000 million between this year and next, far faster than the rate of inflation. The snag is that most of this massive increase of expenditure is earmarked for military research and development. Agencies such as the National Institutes of Health have done far less well, and will be hard pressed to rub along productively with what the Administration now proposes. Only Dr Edward Knapp's National Science Foundation, almost the smallest of the spending agencies, has done reasonably well in this first carve-up, with an increase of 17 per cent (see *Nature* 23/30 December 1982, p.572).

Unfortunately, as the past four years have shown, the President's budget is increasingly a poor guide to what the federal government will spend in the year beginning on 1 October. (Some parts of the budget for the present year are still unsettled.) This year, Congress is if anything more likely than in the recent past to want to have its say. The overall deficit for next year's operations, not far short of \$200,000 million, is not merely an anxiety for the Administration and a challenge to the new Congress but also a possible cause of further worldwide stagnation. Congressmen will be on their mettle to trim the plans for next year's spending. Although the first target seems to be the huge defence budget, much less conspicuous forms of expenditure will not always go unscathed. There is also likely to be strong resistance to the proposals that the pay of public servants should be frozen and social benefits postponed, making penny-pinching respectable wherever it can be practised. Thus the best immediate hope for research and development is that Congress will take to heart what the President was saying last week about the importance of technology in the United States, and that the heads of the research and development agencies will be able to tell eloquent tales about the promise of what they are about. That should not be difficult.

The President's proposals for the improvement of mathematics and science education in the schools are more of a surprise. One of the minor but obvious features of his first two budgets was their meanness towards the educational programmes of the National Science Foundation, the chief source of support for activities such as curriculum development at the high-school level. The question now is not what has changed the President's mind, but who. Is this first blood for Dr George ("Jay") Keyworth, the President's science adviser? Yet it remains to be seen how the job will be done. The National Science Foundation plans to spend a modest amount of money in helping to persuade industry to spend much larger sums on technical education, but what happens in the high schools must depend on the way in which school administrations respond to what the President has been saying — and on the willingness of state governments to pass on whatever funds come their way. And as elsewhere, the problems of teaching mathematics and science well are structural rather than financial. Shortages of able and willing teachers, and the organization of schools, matter more than money. The outlook in higher education is even more clouded. If tax exemption on college fees increases student demand for higher education, the universities will benefit at the margin. But only with the passage of time. In the short run, their best hope is the research budget.

US budget

Generosity and selectivity the passwords to 1984

Washington

HALFWAY through his term of office, and under fire for the poor state of the US economy, President Ronald Reagan has turned to science policy as part of his plan to remedy the nation's economic problems and his own political troubles.

The budget that Reagan submitted to Congress on 31 January reflects this change by offering a large 29 per cent increase in defence research and development, and a 10 per cent increase in basic research in the civilian and military sectors. In some basic research fields, increases as high as 25 per cent are proposed, while some applied work, such as solar and fossil fuels research, is being cut. Presidential science adviser George Keyworth explained that the budget reflects two main goals: first, that of strengthening national defence and second that of providing US industry with an adequate future supply of scientists and engineers trained in US universities.

The President himself signalled this new attention to his science policies with a brief visit to Massachusetts on 26 January. He toured some high technology firms and visited a training centre in the drab community of Roxbury, south of Boston, where the "economically disadvantaged" are taught computer skills backed in part by the IBM Corporation. Afterwards, the President told a group of company executives that he would be an "apostle" for their success, and would help them by increasing basic research funds to universities.

Keyworth explained that the largesse to basic science grew out of a series of cabinet council meetings in recent months run by the President. While arguing about tax policies and other ways of helping US industry, the "lowest common denominator" of agreement in the group was the need for more technically trained personnel. "This thrust came straight from the President", Keyworth told a gathering of reporters last Monday.

The budget proposes an increase in obligations for all federal research and development to rise to \$47,000 million, an increase of 17 per cent over the estimated level for fiscal year 1983. Of this, \$29,900 million will go to defence research and development, an increase of 29 per cent over 1983. The budget for the National Science Foundation (NSF) would be increased by 17 per cent to \$1,200 million.

Although the other agency totals look approximately level (see table), and while the increases for the National Institutes of Health (NIH) are conspicuously low, there

are nonetheless dramatic changes proposed within agency budgets. Keyworth said the Administration had abandoned the previous practice of distributing funding increases evenly among different disciplines, so that while some things — such as basic mathematics funded by NSF, and astronomy — increase dramatically, others are unchanged or decreased.

The Administration's priorities are likely to be changed by the newly elected Congress which will act on this budget this year. Last year, the previous Congress transformed the proposed fiscal year 1983 research budget, and the new Congress is expected to be even less susceptible to presidential influence. So it remains to be seen if the President, having announced a science policy of sorts, will be allowed to implement it by getting his way with the fiscal year 1984 budget.

Among the detailed proposals are the following:

Health

The National Institutes of Health (NIH) budget receive a small \$71 million increase to \$3,771 million — because, Keyworth said, the biomedical sciences are already healthy. But he admitted that Congress is likely to find the sum inadequate and increase it, as has happened in past years. NIH director James Wyngaarden has set several "priority" areas for new funds, but plans for what would be cut are not available. There seems to be some confusion whether the Administration is abandoning the goal of funding 5,000 new investigator grants each year at NIH. One government spokesman said that there could be only 3,676 new grants; but Keyworth said 5,000 could be funded after all.

The other items in the Health and Human Services budget reflect the

President's wish to cut government funded human services: the research activities of offices such as the Office of Human Development Services, are being cut. But an earlier plan to axe the post of Assistant Secretary for Health seems in abeyance, as the budget proposes a \$2 million increase for his office's research and development activities, to \$18 million (see *Nature* 9 December 1982, p.469).

Science

Overall support for NSF research grants is to increase by 17 per cent, to \$1,240 million. Of this, \$180 million will be earmarked for upgrading ageing laboratory equipment at universities, although NSF director Edward Knapp admits that since these funds will be incorporated into research grants, there is no certainty that the full amount will be spent on instruments.

Space

The planetary programme will receive a much needed shot in the arm from the approval of the Venus Radar Mapper mission. This scaled-down version of the now abandoned Venus Orbiting Imaging Radar (VOIR) proposal would get under way in 1984 with a launch in 1988 (see *Nature* 20 January, p.186). Another significant boost for space sciences is the approval of a new Explorer satellite for ultraviolet astronomy.

NASA lost again, however, in its continuing fight to have a fifth shuttle orbiter approved, but development of the Centaur upper stage for the shuttle will continue.

Energy

High-energy and nuclear physics research is to receive a modest increase of \$90 million over the 1983 level to \$553 million. The big winner is the Stanford Linear Collider (SLC) which has received the go-ahead for a 1984 start. This fast construction schedule, which had been pressed for by SLC director Burton Richter, would put SLC in operation in autumn 1986, ahead of its European rival LEP.

Construction of Fermilab's Tevatron will continue, and a National Materials

US budget expenditure on research and development for 1982-84

	1982 Actual	1983 Estimated	1984 Estimated
Defence-military	20,576	23,179	29,882
Energy-related	4,758	4,712	4,713
Health + human services	3,935	4,316	4,416
(NIH)	(3,432)	(3,771)	(3,842)
NASA	3,084	2,506	2,473
NSF	975	1,060	1,240
Agriculture	798	850	849
EPA	335	241	208
All others	1,894	1,996	2,017
R&D facilities	1,234	1,249	1,195
Total	37,588	40,109	46,991

Source: Office of Management and Budget. All figures are in millions of dollars and refer to fiscal years which begin on 1 October of the previous year.

Research Center will be established at Lawrence Berkeley Laboratory, but Brookhaven's ill-fated ISABELLE will remain on ice.

The Administration is equally cool about solar energy, conservation and fossil energy research. It hopes once again to cut these programmes, this time by \$405 million — over half of the current \$707 million budget. If past years are a reliable guide, Congress is likely to restore the cuts.

Magnetic fusion is to be supported at \$467 million, unchanged from 1983. The Clinch River Breeder Reactor is to receive an increase of \$62 million, up from the 1983 level of \$541. This project's popularity in Congress, however, continues to fade.

Defence

Increases in the research budget are overwhelmingly in the defence sector. Most of the \$29,900 million will go for hardware development of big items such as the MX missile. The basic research part of the defence budget would increase by 12.7 per cent to \$867 million, while the budget for the Defense Advanced Research Projects Agency, which funds cutting-edge research and technology, would rise 19 per cent to \$868 million. Areas of emphasis include very high speed integrated circuits, protection against chemical agents, man-machine interface, information processing, and fault-resistant electronics.

Education

NSF is to receive \$20 million for a programme, begun this year on a \$15 million budget, of workshops and training for secondary school science and mathematics teachers. Approximately 10,000 teachers would participate each year. A second programme, administered by the Department of Education through block grants to local school districts, would train 30,000 new science and mathematics teachers over a four-year period. Both programmes would require matching grants, presumably from industry.

NSF is also to receive an increase in its graduate fellowship funding; each fellowship will carry a stipend of \$8,100 a year, up from the current \$6,900. The number of fellowships, now 1,390, would remain the same.

Miscellaneous

The small competitive grants programme in the Department of Agriculture, at present only \$16 million, would be increased by 50 per cent. Grants would be offered for the first time for research in the animal sciences. The increase is also to take into account the new agricultural research opportunities in genetic engineering.

The National Bureau of Standards faces a \$16 million cut in research and development funding. Support for the Environmental Protection Agency's research programme would be cut 14 per cent from 1983 levels, to \$208 million.

Deborah Shapley & Stephen Budiansky

India in Antarctica

International treaty still on ice

Washington

INDIA'S Antarctic ambitions, hitherto not clearly explained, owe something to its wish to be the leader of developing countries seeking a voice in Antarctic affairs. This was the impression left last week by the visit of Dr S. Z. Quazim, secretary of India's Department of Ocean Developments.

Whether developing countries will have a say in the future of the Antarctic region, and of the 1961 Antarctic Treaty that governs it, is not, however, clear. In recent years, some developing nations have started to work under the framework of the treaty, which is subscribed to mainly by industrial nations rich enough to send ships, build bases and carry out scientific programmes in the Antarctic. Other developing nations have denounced the treaty arrangement, saying the United Nations should control Antarctica. India has not yet taken sides on this question although several treaty nations, including the Soviet Union, are urging India to join. Quazim said the Indian government is studying the advantages of acceding to the treaty, which freezes the dispute over territorial claims and bans new claims and military activities.

India is the only nation to have sent major expeditions to Antarctica without first acceding to the treaty. Quazim led the first Indian expedition in 1981–82. A second expedition is there now, preparing for a permanent base on the Princess Astrid Coast, in a part of the continent claimed by Norway. The site is said to be at 70° 46' S and 11° 50' E, about 50 miles from the Soviet station Novolazarevskaya. India also has an unmanned weather station at approximately 43° E on the coast there.

Technically, the treaty nations already in Antarctica can refuse to help the Indian expedition, because it is operating outside the framework of the treaty. Technically also, India is free to assert a territorial claim only so long as it stays outside the treaty. But nobody seems to be considering either of these extreme actions.

Francis Johnson, the assistant director of the National Science Foundation, which runs the US Antarctic Program and who had a meeting with Quazim, says "if something realistic emerged on which we could cooperate, then the policy question would have to be faced. But we do not intend to enter into cooperative activities with any nation that has not acceded to the Antarctic Treaty."

One sure sign of India's continuing plans for the Antarctic is that Indian officials in Washington are negotiating to buy as many as five ski-equipped Lockheed C-130s. Lockheed is said to have sought permission from the State Department to export the planes. The United States operates six C-130s in Antarctica each season, and at current prices a single aircraft, with spare parts, costs around \$26 million.

The ultimate question is whether India will accede to the treaty rather than press for United Nations control. At the time of the 1981–82 expedition, Quazim said, India sent letters to other non-aligned nations describing its interests in the Antarctic. Few raised objections.

On the other hand some Indian scientists have attended a recent meeting of the Scientific Committee on Antarctic Research (SCAR) held in Leningrad, Quazim said. Both Brazilian and Chinese scientists have participated in several SCAR meetings, and have visited US bases in the Antarctic. Both of these influential developing countries have indicated their intention of working through the treaty framework. Active participation in SCAR is considered a prerequisite for becoming a full party to the Antarctic Treaty.

US officials are wary of getting involved in India's plans in the Antarctic so long as India stays outside the treaty. But having a major non-aligned nation inside the treaty circle might be workable, one official said. "It would be like having a socialist who always runs for President. The socialist never wins, but he has some influence." Clearly, some Indians are thinking the same thing.

Deborah Shapley



The ubiquitous Lockheed C-130 on skis — India would like five

Science in Greece

Starting almost from scratch

GREECE is about to take political notice of science — which means it may have a science policy, and a reasonable science budget, for the first time since the sack of Alexandria. According to the newly-appointed minister of research and technology, Professor George Lianis, the Greek government aims to multiply its research and development spending sixfold over the next five years — and to reform or create whatever institutions are necessary in the process.

Lianis knows about reform in Greece: he was one of a team of radical ministers appointed to the ministry of education by socialist Prime Minister Andreas Papandreou, when the PASOK party came to power in October 1981. That team attempted to bring the universities up to date, and in less than a year pushed through a bill which reduced the power of the previously autocratic professors, gave recognition and a career structure to other members of staff, established graduate schools and offered a degree of democratic power to the students.

Those major reforms are now working themselves through, and Lianis — an engineer trained at Imperial College, London — has turned to science. Melina Mercouri, the former actress, appointed minister of culture and science, will remain in that role, but her primary preoccupation will be the arts and archaeology.

Greek science is both “stifled by bureaucracy” and “anarchic” says Lianis, one of whose priorities will be to sort out which scientists in the universities and the few research institutions (such as the Hellenic Research Foundation and the Democritus nuclear research centre) are more or less political appointees of previous governments, and which are doing real science.

Then there must be proper structures for scientific assessment, and the allocation of grants. According to one Greek scientist on the telephone from Athens last week, “no-

one has been properly reviewed before”. For researchers outside the university reform, there is at present neither a career structure nor tenure, but a bill is being prepared to correct these failings.

Lianis hopes for much from expatriate Greek scientists — like himself — whom he expects will return out of a sense of duty, “nostalgia” for the home country, recognition and status, and an opportunity to create something new. “People will have to come and create the conditions for good research in Greece” he says.

It is early to ask for policy, but Lianis stresses software and informatics, biotechnology, solar and geothermal energy, and the life sciences (for health and agriculture). He has been particularly impressed by a study which showed that Greece, with its many islands, fishing ports

and warm waters, has tremendous potential for aquaculture, and he is seeking cooperation with Italy in this field.

Lianis's budget this year will be 3,000 million drachmas (£21 million), including the salaries of his staff of 70 and the staff of the research institute — an apparently tiny sum, but in terms of equipping and establishing laboratories already double what was available last year. More cash is available to the universities and research institutes by direct funding of special projects through the treasury.

Lianis wishes to push the total Greek expenditure on research and development from 0.2 per cent of gross national product (1982) to 1.2 per cent (by 1987). Some of that, he hopes, will come from industry, which at present spends “next-to-nothing”.

Robert Walgate

Heavy ion research

More machines come on-line

THE French Grand Accélérateur National d'Ions Lourds (GANIL), a FF 557 million (£52.5 million) heavy ion accelerator designed to probe a new energy region for nuclei, is officially opened this week by the French research minister. But GANIL has already produced its first remarkable results which throw a new light on the physics it will be surveying.

These results appear to show that nuclei “fragment” and behave like collections of individual particles at GANIL energies of around 45 MeV per nucleon — a phenomenon hitherto believed to set in only at the much higher energies of 1,800 MeV per nucleon used by the Bevalac heavy ion machine at Berkeley, California. If these early results of an Orsay group working at GANIL are confirmed, the “intermediate energy phenomena” which were to have been studied may not in fact exist.

The original argument went like this: GANIL would be aimed at an unexplored energy region between the 18 MeV per nucleon available at GSI Darmstadt in West Germany (and at Dubna in the Soviet Union), and the 1,800 MeV per nucleon available at the Bevalac machine. At the former energies, collisions yield high-mass composite nuclei, often at high states of spin; at the latter, nuclei appear to behave like gases of protons, neutrons and other nuclear fragments. In between, GANIL hoped — and hopes — to study a region where nuclei behave like collective nuclear “matter”, a material independent of particular nuclear structures or shapes but still composite: broadly speaking, the liquid state rather than the gas visible at Berkeley.

However, the first GANIL results appear far more like Berkeley measurements than expected: in particular, the projectile nucleus (argon-40 at 45 MeV per nucleon) fragments on hitting its target into a spectrum of smaller nuclei (including phosphorus, chlorine and sulphur) which

fly on with the same velocity as the projectile.

Meanwhile in Britain, two other accelerators have come on-line. One of them, the Nuclear Structure Facility, a 20 MV Van de Graaff electrostatic heavy ion accelerator at Daresbury, has produced its first results, and the other, the Spallation Neutron Source (for condensed matter studies) at the Rutherford and Appleton Laboratory at Didcot its first test beam.

At the Nuclear Structure Facility, a facility which works at lower energy than GANIL, and which cost £13.6 million and was eight years in construction, the first experiments — a collaboration between groups from the universities of Liverpool, Manchester and Copenhagen — have investigated high nuclear spin states using hafnium-168 and titanium-48 beams, and discovered clear evidence for “spin depairing”. Group spokesman John Lisle of the University of Manchester describes it thus: at low energies, the nucleons in nuclei have paired spins, and the nuclear material behaves rather like a superconductor (in which electrons pair up); but at high energies the nuclear spins appear to uncouple, producing something like a normal metal. Thus the normal state of nuclei appears like the abnormal (superconducting) state of metals, and the abnormal (excited) state of nuclei like the normal state of metals. Sketchy evidence for this was available a year ago, said Dr Lisle, but the Nuclear Structure Facility results show it clearly for the first time.

The Nuclear Structure Facility appears to have had this success not just because of the accelerator — a close equivalent at Oak Ridge National Laboratory in the United States has been running for a year, and another is ready in Japan — but because of a unique gamma-ray apparatus called TESSA, which was used to analyse the decay spectra of the spinning nuclei with great accuracy.

Robert Walgate



George Lianis heads a new ministry

El Salvador physicians

US medical team reports abuses

Washington

HARASSMENT, exile and "disappearances" of doctors in El Salvador have led to an overwhelming deterioration in the health care system of that country, according to a delegation of US doctors and nurses that has just returned from there.

The delegation, sponsored by the American Public Health Association and several other US medical organizations, found a continuation of the threats, harassment and killings that have driven many of the ablest doctors from the country. Eight of the 11 doctors who formed an organization (the National Committee for the Defense of Patients, Workers and Health Institutions) to oppose armed attacks on hospitals have disappeared or been killed since 1980; workers for the Green Cross (a volunteer organization that provides relief services) have been arrested and tortured and, in some cases, have disappeared or been found murdered, the delegation said.

"To organize in El Salvador is to commit a dangerous act", said Dr John Stanbury



of Massachusetts General Hospital, a member of the delegation — "no matter what the organization is." He said the physicians' organizations did not appear to have any political cast. Dr David Halperin, a surgeon at the Kennebec Medical Center in Maine, added that the "crime" of "physician after physician we talked to in Mexico who had been forced to leave [El Salvador] . . . was organizing to provide better medical care".

Peter Romero of the US State Department's El Salvador desk, who talked to the delegation upon their return, said: "I agree with them that the health situation is deplorable. But their allegation that the government is responsible is manifestly unfair."

Romero said that in the eastern part of the country alone, 21 health clinics had been destroyed by guerrillas. And while conceding that "there are far too many people who are being incarcerated for

nothing more than affiliation". Romero said that "many of these organizations belong to the guerrilla Left". He declined to answer questions about the political allegiance of specific medical organizations.

The Reagan Administration last week decided to continue arms supplies to El Salvador, citing as justification some improvements in human rights. But Halperin suggested that some of these improvements may actually be signs of deterioration. For example, harassment of doctors had declined "because they are no longer trying to organize". Raids on hospitals by military and paramilitary right-wing groups, some of whom in 1980 shot suspected guerrillas on the operating table, had stopped because guerrillas are no longer being treated in public hospitals, he said.

The country's one medical school, seized by government forces in 1980, remains closed, Halperin said, its library burned, laboratories destroyed and equipment looted.

The delegation also visited a camp in San Salvador for refugees from the "contested" part of the country, and found "appalling" health conditions, including widespread malnutrition. Political prisoners at two major prisons that the delegation was allowed to visit showed physical signs of torture, including beatings and chemical burns.

Romero at the State Department said the United States would look into two specific cases for which the delegation provided evidence of torture. But he said the delegation "was not able to tell me that a majority or even a small minority [of the prisoners] were being tortured". Romero added that the doctors had "not really" told him anything new about conditions in El Salvador.

Asked about Romero's comment, Halperin said, "Maybe it wasn't news to him, but every single woman prisoner we interviewed said they had been sexually abused. Every single male prisoner had been subjected to mock executions."

Halperin said a small sign of encouragement was the government's recent decision to allow one doctor and one nurse from the International Committee of the Red Cross to travel freely in the eastern part of the country. The Red Cross has also set up a blood-banking system in San Salvador, helping to ease the serious blood supply problems that a 1980 delegation of US doctors found.

But Stanbury said that, overall, "everybody we talked to said they thought things were worse", a view that members of the delegation will present to Congress this month at hearings on the administration's certification of El Salvador's progress on human rights.

Stephen Budiansky

US computer research

Inman to join cooperative

Washington

THE appointment of Admiral "Bobby" R. Inman as president and chief executive may be the best thing yet to have happened to Microelectronics and Computer Technology Corporation, the research and development organization set up last August by ten computer companies. Inman, a career naval officer for 31 years who has been in his time director of the National Security Agency and deputy director of the Central Intelligence Agency, made waves last year by telling the academic community that self-restraint was the best way of avoiding government restrictions on the publication of sensitive research (see *Nature* 22 April 1982, p.693).

Admiral Inman says that after leaving the Navy last fall, he had been negotiating a post with a major corporation, but that he had been attracted to the proposal that he should run the computer cooperative by the challenge of "trying to get US companies to pull together" in a fast-moving field where the outcome could determine "economic success or failure for the United States".

The new corporation, known as MCT, has also just been told by the Justice Department that collaboration on research will not offend anti-trust legislation, although that ruling could yet be challenged by corporations outside the existing consortium, which includes Advanced Micro Devices Inc., Control Data Corporation, Digital Equipment Corporation, Harris Corporation, Honeywell Inc., Motorola Inc., NCR Corporation, National Semiconductor Corporation, RCA Corporation and Sperry Corporation. A spokesman for IBM — with AT&T a conspicuous absentee from the list — said last week that his corporation's decision not to join had been reached on the basis that membership would not be in the best interest of its business. MCT is, however, still negotiating with two dozen other companies.

The organization of MCT should be clearer after the first board meeting arranged for 4 February, when guidelines for participation will be formally adopted. The present intention is that the members will contribute separately, in proportions for the time being confidential, to the cost of one or more of the four research programmes, and will be free to exploit the results independently as these are made available. There is also a scheme by which corporations and universities may become associate members.

The next step for MCT will be to recruit project leaders for the four agreed research programmes in advanced computer architecture, computer-aided design and

manufacture based on high-density semiconductor devices, software productivity and packaging — the optimal packing of microcircuits on semiconductor chips.

Similarly, the scale of operations has not yet been decided, although an annual budget of more than \$50 million was talked about last August. Admiral Inman said on the telephone last week, however, that he had received assurances from the member companies of their commitment to the enterprise, and that the next few months should show how well they deliver.

Inman's special value to the enterprise will be his experience in the development of large computer systems for the National Security Agency and his reputation for leadership. But MCT says that he has not been hired as a "government contract-broker" although there is a possibility that the company may qualify for research and development grants in its field. □

PWR league table

Germany tops

GERMAN pressurized water reactors (PWRs) are the best in the world, except for isolated stations in small countries such as Switzerland, according to a UK study by the Science Policy Research Unit (SPRU) at the University of Sussex. The study, compiled by SPRU researcher Steve Thomas, compares plant availability (actual energy generated divided by theoretical maximum) for 1978–81.

West Germany, with six operating PWRs in 1981, turned in an average availability of 79 per cent that year — compared with 57 per cent in the United States and in Japan. However, Germany's six boiling water reactors (BWRs) did much worse: 43 per cent in 1981. French PWRs in 1981 fell mid-range at 67 per cent.

Among suppliers, Westinghouse did best with its two-loop (500–700 MW) PWRs: 75 per cent. But its three- and four-loop PWRs (over 1,000 MW, like Britain's proposed Sizewell design) showed consistently less than 60 per cent. The key problem was steam generator failure, which the German modifications of the Westinghouse design appear to have solved. A second supplier, Babcock & Wilcox, "have a poor record" even apart from the Three Mile Island accident.

Among reactor types, the unchallenged lead is held by the Canadian heavy water reactors (CANDU), which were precision engineered to save costly heavy water. The eight CANDUs in operation in Canada showed availabilities of nearly 90 per cent in 1981, and one in Argentina 94 per cent.

These figures raise new issues of competition for reactor vendors in the world market, says Thomas. In the previous period of large ordering (1967–73) utilities based their orders on promise of performance. Now they have real figures on which to make their judgements.

Robert Walgate

Consent in clinical trials

Will verbal assent suffice?

RULES for obtaining patients' consent to participate in clinical trials at one of Britain's leading teaching hospitals will soon be altered if the institution's ethical committee gets its way. But, according to the *Acton Gazette* last week, the hospital's governing body is resisting a proposal to replace verbal by written consent.

The Hammersmith Hospital's Research Ethics Committee had suggested some months ago that present procedures for obtaining consent, which involve the patient reading and signing a stark clinical description of his or her condition, might in certain cases cause unnecessary distress. It recommended that when consent is being sought from patients with cancer or malignant blood disease, a "sensitive" verbal explanation of the proposals should be used instead, and that the patient's oral approval, witnessed by a disinterested third party, should suffice as evidence of consent.

The proposal, later modified so that a written copy of the oral explanation would also be made available to the patient, has now been given a cool reception by the Hammersmith Special Health Authority, which apparently felt that if a written copy was to be made available, the patient might as well be asked to sign it as confirmatory evidence of consent. The proposal has been referred back to the ethical committee, to the annoyance of some members of the medical school.

Legally, informed consent has to be obtained from any patient participating in a clinical trial, and the British Medical

Association's Handbook of Ethics stresses that "as full an explanation as possible" should be given to potential subjects. Guidelines issued by the Medical Research Council require "consent freely given with a proper understanding of what is proposed" for clinical trials. But the consent does not have to be in writing, although this is usually considered desirable.

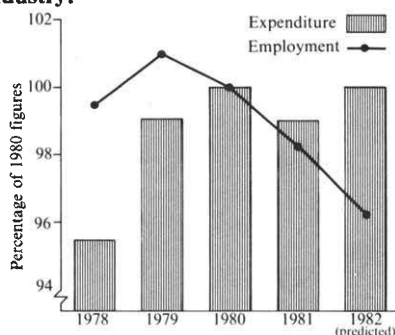
So have the committee's proposals come full circle? Not quite. If the ethics committee is able to satisfy the health authority's criticisms, the patients concerned will first receive a verbal description of the proposed trial from their own doctor which will be couched in "appropriate terminology". The written consent requested by the health authority will be based on that description rather than on a bald medical one.

Professor John Goodwin, chairman of the Research Ethics Committee, says that the proposals would not significantly alter the number of patients participating in clinical trials (most of those asked agree anyway) but that they would give seriously ill patients more peace of mind. But some still have doubts. Special Health Authority chairman Mr Christopher Bland said: "We fully appreciate that the motive behind the proposal was to spare terminal cancer patients the possible mental suffering caused by very stark written explanations. But at present the absolute clarity of understanding required before getting informed consent to taking part in clinical research might not be achieved."

Tim Beardsley & Martin Flanagan

UK industrial research expenditure and jobs

WHILST employment in the research and development sector of British industry has fallen since 1978, expenditure has merely levelled out, according to figures published by the Confederation of British Industry (see figure). But there are sharp differences between individual industries. Thus spending has increased in both the chemical and electronic industries but has been markedly reduced in the automobile industry.



Industrial expenditure on research and development

Sector	1978	1979	1980	1981	1982
Vehicles (incl. aerospace)	94	97	100	94	-
Chemicals (incl. petroleum products)	98	98	100	101	-
Electronics	93	102	100	107	-
Food, drink, tobacco	99	97	100	99	-
Total	95	99	100	99(100)	-

Source: R & D — in recession too?; CBI, London, 1983. The 1980 value is taken as 100 per cent.

The CBI report finds little encouragement from comparison of Britain's present spending on research with that of other industrialized countries. After only weak growth during the 1960s and 1970s, the British research budget remains less than half that of West Germany and still lags well behind that of France. Melanie Kee

Australia in space

New consortium blasts off

Canberra

AUSTRALIA's first space company, Auspace Pty Ltd, created on 5 January by collaboration between SA Matra, the French aerospace company, and Hawker de Havilland (Aust) Pty Ltd, has already tendered successfully for Starlab, the projected joint US-Canadian-Australian space telescope. A subsidiary of the British Hawker group (now part of British Aerospace), Hawker de Havilland owns 60 per cent of the new company and has no Australian equity. Matra, which owns the other 40 per cent, has been involved with Starlab for about two years, supplying it with expertise in space technology and in return accruing government offset credits.

Australian involvement with Starlab, originally intended as a complement to the US National Aeronautics and Space Administration (NASA)'s Space Telescope (ST), dates from an informal suggestion in 1979 by Professor Don Mathewson, director of Mount Stromlo and Siding Spring Observatories (MSSO), after the project was dropped from the NASA budget. Canada joined in shortly afterwards and, as a result of discussions between the three relevant government agencies, Starlab was planned as a free-flight telescope rather than a facility operated in the payload bay of the space shuttle during sorties. Canada will construct the body with the telescope, Australia the instrument package consisting of a camera, spectrograph and detector system, while NASA will develop the space platform, bear the costs of the first two flights by the space shuttle and then operate Starlab in orbit. The tentative launch date is 1990.

Starlab will be a modified Ritchey-Chretien telescope with a primary mirror of 1 metre diameter (ST has a 2.4-metre mirror), with a field of view a hundred times wider than that of ST but with a spatial resolution close to the ST threshold. Starlab will, like ST, accumulate signals from single photons within the wavelength range 100–1,000 nm. Starlab information from a large area of the sky will be used to point the more sensitive ST much as ground-based Schmidt cameras are used to direct large ground-based cameras. Starlab is designed for a lifetime of 20 years, with periodic refurbishment by the space shuttle.

Australia's contribution will be a wide-field camera called the Direct Imager, together with a long slit echelle spectrograph and associated detector systems. The large format photon-counting detectors, with more than 8×10^7 pixels (picture elements), present the greatest challenge to MSSO ingenuity.

For the time being, the MSSO group is engaged on design studies as a "public interest" project almost as an afterthought

with funds left over from last year's Industrial Research and Development Incentives budget. Much of the A\$3.337 million (£2 million) allocated for this phase will flow to the many companies subcontracted by Auspace. If the three partners decide in 1984 that the telescope should be built and launched, the final cost to the government will be about A\$30 million.

Auspace has ambitions beyond Starlab and will look further afield for future contracts. Stanley Schaezel, one of its directors, says that the company will start bidding for international satellite contracts, but not in the European arena against its parent companies.

Vimala Sarma

Todaro to leave NCI

Washington

DR George Todaro, director of the National Cancer Institute (NCI)'s Laboratory of Viral Carcinogenesis is to become scientific director of a new private company, Oncogen, in Seattle, Washington state. The company is a joint venture of Genetic Systems Inc. of Seattle and Syntex Corporation of Palo Alto, California. He will also be professor of pathology at the University of Washington Medical School.

Todaro's associates at his NCI laboratory at the Frederick Cancer Research Facility in Maryland say his departure is unrelated to the recent difficulties the facility has experienced (*Nature* 20 January, p.185). He decided to leave because this opportunity was too good to pass by, as it will allow him to continue his basic research and carry it through to applications for cancer diagnosis and new modes of treatment of cancer.

Todaro plans that the company will carry out research on monoclonal antibodies for the detection and treatment of various human tumours, on viral and cellular oncogenes, on human cancer-causing viruses, and on the role of the growth regulatory peptides in the control of cancer, the focus of much of his current research.

Genetic Systems Inc., which is already established in the building that Oncogen will occupy, is an established biotechnology company listed in *Nature's* biotechnology stock index. It is developing monoclonal antibodies as diagnostic reagents particularly for venereal diseases. Backed in part by Syntex, the company also has ties to the University of Washington Medical School and the Fred Hutchinson Cancer Center in Seattle. Last year Todaro explored the possibility of becoming Director of Scientific Programs at the latter.

Deborah Shapley

Soviet Jewry

Higher hurdles to jump over

BORIS Kanevskii, the Moscow mathematician who since 1979 has monitored the mathematics entrance examinations of Moscow University for anti-semitic bias, was last week sentenced to five years' exile in Siberia. His collaborator Valerii Senderov, in detention pending trial, is a former associate of the independent trade union movement SMOT and may face a heavier sentence. An 18-year-old student, Ilya Geltser, who had helped coordinate the findings, has been given a three-year suspended sentence.

Although anti-semitism is strictly forbidden by the Soviet constitution, during the past ten years a bias against Jews has become increasingly apparent in Soviet academic life. Prejudice is particularly prevalent among mathematicians, reflecting, apparently, the views of a relatively small but influential clique.

The first major *samizdat* account of anti-semitism in Soviet mathematics was compiled in 1976 by Dr Grigorii Freiman, then a professor at the Kalinin State University and a Communist Party member in good standing. He was angered by the repeated refusal of the Higher Attestation Committee to accept the Candidature (PhD) thesis of his best student. Dr Frieman established that in the oral entrance examination in mathematics in particular, applicants with names suggesting Jewish (and also Armenian) origin were hived off into special groups which were set questions beyond the competence of even the brightest school-leavers.

Kanevskii and Senderov set out to provide a statistical comparison of the examination results of Jews and non-Jews in the entrance examinations of 1979–82 at three leading mathematical institutions — Moscow State University, Moscow Institute of Physics and Technology and Moscow Institute of Engineering and Physics. They found that whereas Gentiles (and, in particular, Russians) regularly obtained examination results commensurate with their school performance, qualified Jewish candidates often failed to score pass marks.

Having found that the three leading Moscow mathematical institutes investigated admitted not more than 0.8 per cent of students having at least one Jewish grandparent, Senderov and Kanevskii produced their own stop-gap solution — an unofficial "People's University of the Mathematical Sciences" for Jewish would-be undergraduates who had failed to gain a university place. Although a number of postgraduate seminars had been operating for years among the Moscow Jewish community, this was the first real attempt to do anything at the undergraduate level.

Vera Rich

US medical education

The return of the intern?

Washington

THE American Medical Association (AMA), which represents roughly half of the 470,000 doctors in the United States, has just weighed in with a plea for a return to basics in premedical education and the four years of medical school that, in the United States, leads to the MD degree. In particular, AMA urges a return to the year of internship that was traditionally sandwiched between medical school and postgraduate residency, but which was eliminated eight years ago.

The report, adopted by the AMA House of Delegates last June but only published three weeks ago, reflects the fears of practising physicians that specialization has become so rampant that US medical education now turns out specialists who literally cannot undertake general patient care. "Physicians should be broadly educated if medicine is to continue to be a highly respected profession", it says.

The debate over the internship year has been going on for some time, although the use of the term was officially dropped in the early 1970s, when a shortage of doctors was forecast and when the convention was abandoned that a student should attend medical college and practise internship and residency at three different institutions. A separate internship year has not been required since 1975.

The new report focuses on this year, which is now the first year of the residency, and urges that it be "a broad year of general training" in which students practise in some field other than that in which they plan to specialize. Control of the curriculum should be in the hands of the school itself — not the individual departments. This would presumably enable even the most specialized doctor produced by the system to be able to do general medicine as well.

Some experts in medical education not associated with AMA characterized the report as showing a certain "nostalgia" for a simpler past that many doctors feel. "Specialization is not going away", he said, "not with the explosion in basic science and diagnostic technique that is going on."

The report does not address some other issues that together comprise something of a crisis for medical education. Thus, for the first time in history, the system is producing more MDs than there are residency positions at teaching hospitals, which are shrinking because of cuts in federal funds and other factors.

Indeed, by 1990, there may be 70,000 too many physicians for the population's needs, according to the graduate Medical Education National Advisory Committee. Fifteen specialized fields may have an oversupply, the committee forecasts, including neuroscience, pulmonary

medicine, endocrinology and cardiology. Moreover, the oversupply is expected to make doctors guard their turf, turning over fewer of their responsibilities to cheaper professionals such as nurses — thus driving up the cost of health care. The committee also foresees that as the US population ages and as acute illness declines, there will be relatively less demand for the kind of intensive acute intervention stressed in US teaching hospitals today.

The AMA report is also concerned with the dehumanizing aspects of today's medical school curriculum. "The acquisition of a broad, liberal education leads to such understanding [of the role of medicine in society] and further provides for more effective patient communication, as well as an understanding of the influence of social and economic problems on disease and convalescence."

The report says pre-professional college education has become too technological at the expense of a liberal education that could contribute significantly to rational and ethical judgement, and asks that medical schools let it be known that they seek not only students who have spent their college years cramming pre-med courses but those who are strong in the arts and humanities as well. Faculty evaluations should place more emphasis on a student's integrity and other personal qualities, it says.

The report also says that research funds have not always been a good thing for medical education, and that schools have also been distracted in trying to meet social needs such as the energy problem. Medical schools should not assume "obligations" that compromise "the unique purpose of educating students to become physicians".

There are two other major studies of medical education under way. One is by the Institute of Medicine of the National Research Council, which is making a one-year planning study of medical education, the manpower question in particular. The Association of American Medical Colleges (AAMC) is taking a different approach by trying to involve the colleges themselves in debating their problems. Hearings will be held for the AAMC study in San Francisco, Houston, Chicago and New York. August Swanson, AAMC's director of the department of academic affairs, rejects the complaint that there is too much science in medical education.

Swanson argues that today's students will, as qualified physicians, have repeatedly to adapt to a deepening knowledge of medicine and a changing technology. He argues also that future physicians need an intellectual framework to accommodate new knowledge and developments, not the rote learning with which they are now provided.

Deborah Shapley

West German research

Time to count costs

Now that the West German economy is belatedly suffering the effects of recession, questions are being asked as to whether the money invested in research is being well spent and how the achievements compare with those in other countries. As a contribution to this debate, the Deutsche Forschungsgemeinschaft (DFG), which distributes most of the public money for pure research, has commissioned reviews from about 100 researchers in universities, Max Planck Institutes and industry on the state of research in their own fields. Contributors were also asked to identify conditions needed for successful research and comment on the extent to which these are fulfilled in West Germany. The reviews have now been published in a single volume, edited by Dr Cristoph Schneider*.

The reviews paint a picture of rich variety supported by generally adequate funding. DFG, the Max Planck Gesellschaft and the Federal Ministry for Research and Technology are commended for providing their funds with the minimum of strings. In the Max Planck Institutes, where most of the money from the Max Planck Gesellschaft is spent, bureaucratic barriers have not been a brake in the past, although they are said to be increasing.

A more worrying picture emerges from the universities. Here the heavy load of administration, teaching and dealing with student affairs leaves insufficient time for effective research. There are also major problems for those in the 30–40 age group who have few prospects of a teaching post because the present holders will not retire much before the end of the century.

Many of the researchers in reviewing their work take the opportunity of stressing that research can often be slow to yield tangible results. The audience being addressed here is clearly that of politicians and accountants, who with their eyes on the next election or the bottom line in the annual report, are not always prepared to support long-term projects.

Perhaps the long haul involved in research is part of the explanation of the small number of Nobel Prizes won by Germans since the war. It was not until after 1960, when the disasters of the decade before 1945 had been overcome that a new generation of graduates entered research and funding was increased. Since the quality of facilities in West Germany appears now to match the best, it can be expected that Germany will move up the league in the next 20 years, provided that its scientists can overcome their alleged fear of risk taking.

J.S. Dunnett

**Forschung in der Bundesrepublik Deutschland — Beispiele, Kritik, Vorschläge* (ed. Schneider, C.) (Verlag Chemie, Weinheim, 1983).

Wegener polar institute

SIR — On page 267 of *Nature* of 27 May 1982, the function and significance of the "Wissenschaftsrat" in the federal system of West Germany is described. There is, however, a serious mistake as far as polar research is concerned.

The article states: "the council was . . . influential in the decision that the new Alfred Wegener research centre for polar research should be located at Bremen rather than at Kiel". This is wrong. The comments of the Wissenschaftsrat itself on the question of where to locate the polar institute, as well as many reports in all German newspapers, show that the Wissenschaftsrat favoured Kiel as the site for the new polar research institute.

The reasons for the decision of the Wissenschaftsrat was that Kiel had been the centre of oceanography for more than 100 years. The Institute of Oceanography at the University of Kiel ranks among the most important research centres of its kind in Europe. Kiel University disposes of a considerable amount of relevant research potential in other scientific institutes which have for long been performing oceanographic research work together with our Institute of Oceanography on nearly every ocean in the world.

The decision of the then government of West Germany in favour of the polar research institute being located at Bremen was definitely taken against the recommendation of the Wissenschaftsrat and has to be judged according to criteria other than scientific ones.

V. BAGAN

Präsidium der Universität Kiel,
Kiel, FRG

Future of universities

SIR — As the registrar of the university forced to think most radically about its future developments and purposes, I was naturally interested in your article on the future agenda for the next chairman of the University Grants Committee (UGC) (*Nature* 300, 301–302; 1982).

Not all British universities seek to imitate Oxbridge. The very first annual report of the first Vice-Chancellor of the University of Salford stated that "in its concern to be seen as equal [to other universities], the university must not accept that all the virtues rest in Oxbridge or in the great provincial universities or in the new foundations of the early '60s. Salford must not become a pale shadow of what others already are".

My university has for a considerable time been seeking to have accepted a different ideal — and not that of Imperial College, excellent though that institution is seen to be. The truth as you clearly imply later in your article is that universities are judged,

except for a very, very few cases, by the Oxbridge intellectual and academic model — the most grievous mistake that those in charge of the university system could have made. But the general coherence and thrust of your otherwise trenchant plea for innovation and well-directed change may be undermined by a less than scrupulous regard for the facts.

I really do think that you have distorted almost beyond recognition the relationship between UCCA and the universities. My own salary may be paid with the aid of a computer. It is not, however, (I trust) the chance product of some random process, any more than is the computer-generated letter of offer or indeed rejection which is sent to candidates for admission to universities. University selectors are firmly in control of the choice of students: and it does you no credit to promulgate innuendoes to contrary effect.

Your arguments since mid-1979 for institutional diversity, for incentives and for innovation have been a powerful indictment of the way in which the university system, with the connivance of the government, has been allowed to ignore the real question which the government — belatedly and for all the wrong reasons — set both itself, the UGC and the universities in 1981. There is as yet no sign of any real change in the underlying consensus — and complacency — on which the UGC gambled so heavily in July 1981.

S.R. BOSWORTH

University of Salford, UK

Friend or FoE?

SIR — I wish to protest in the strongest possible terms at your abuse of editorial privilege in respect of the two editorials "How to get off the MX hook" and "Inventing the wheel" in *Nature* of 23/30 December 1982 (pages 671–672).

On the one hand anyone who can write "The ideal . . . would be that missiles such as MX should be made . . . invulnerable . . ." has no conception of the true meaning of the word ideal. In an ideal world there would be no missiles, MX or otherwise. The manufacture and disposition of any such offensive weapons is a political issue — not a scientific one — and any discussion of these matters has no place in a scientific journal.

Again I would have thought that the editor of a journal so quaintly named as *Nature* would have hesitated to criticize the "Friends of the Earth" in this respect. Nor is it proper to deride their views editorially. Whilst the engineering design of these reactors may be above reproach, which I doubt, their construction in the United Kingdom is certainly open to criticism on both economic and ecological grounds.

D.C. PRESSEY

Rickmansworth, Herts, UK

Earthquake light phenomena

SIR — During the General Assembly of the European Seismological Commission and the meeting of the Subcommittee on Earthquake Prediction Research at Leeds, on 27 August 1982, the "Project on Collection and Evaluation of Data on Earthquake Light Phenomena" was initiated.

The true nature of earthquake lights (EQL) remains obscure. Thousands of reports on them are available, and a few drawings and even photographs have been obtained, but the origin of EQLs is still unknown. Apart from studies in the 1930s and 1940s, no systematic study of an EQL has been made.

The first objective of the new project is the collection and analysis of old (historic) and recent EQL-data from all over the world. Particularly important are the relationships between the occurrence of EQL and the magnitude (and focal depth) of the earthquake.

In many cases EQL have appeared before impending shocks, not only during and/or after them. Hence the second aim of the project, the rapid publication of any observations, so that the data are available for all those dealing with earthquake prediction.

Readers are invited to send reports of any possible EQL phenomena to me at the address given below.

P. HÉDERVARI

Georgiana Observatory,
H 1023 Budapest,
Arpád fejedelem útja 40–41, Hungary

Seeing stars

SIR — Why single out the BBC's breakfast-time astrologer for a public upbraiding? (*Opinion*, *Nature* 20 January p.184). Why not go the whole way and include, inter alia, politicians, union leaders, captains of industry, Bobby Robson, and (worst of all) economists? My licence fee is continually being dissipated to allow all of them to make the most incredibly inaccurate predictions. I couldn't take the harsh reality of it all until, rather than turning to alcohol, I began to see the funny side of things. I was even (dare I say it) laughing occasionally! But now I read words like "hoodwinked", "public dis-service", "pernicious". Now I'm hopelessly lost, can anybody help me?

Sorry, I think the Great British Public (GBP) is a little more discerning than you would like to think. Even *Nature* readers know when things are being taken a little bit too seriously.

FRANK H. ALLEN

University Chemical Laboratory,
Cambridge, UK

Oncogenes, reductionism and all that

Much of the excitement generated of late by studies of human oncogenes stems from the belief that the causation of cancer will be better understood. But is that not naive reductionism? Simply: NO

THE excitement in the past few months about the oncogenes, especially that generated by the discovery (see *Nature* 300, 143 and 149; 1982) that a single point mutation in a gene coding for a protein called p21 appears to characterize some human carcinomas, has now predictably begun to stimulate scepticism. Naturally, in such a novel field of enquiry, there is a host of technical questions to be asked. Is the occurrence of the oncogenes now thought to be associated with malignancy confined to overtly malignant cells? And to what extent do the special properties of the NIH 3T3 cells so far used as an assay of malignancy predetermine some of the results so far obtained? Such questions are plainly important, but the answers in due course provided for them are unlikely to diminish the central importance of the recognition that there is, at last, an identifiable material, the protein p21, whose structure appears to be intimately involved in the difference between normalcy and malignancy. But can a biological phenomenon as complicated as cancer be so simply explained? Is it not outrageously reductionist to suppose that a molecular account can be given of a biological aberration known to involve many kinds of cellular interactions as well as environmental and even psychological influences?

Reductionism

That question has not yet been fully debated in *Nature*. No doubt, in due course, it will be. The question was, however, raised the other day by a correspondent who had also written similarly to another journal. (Ironically, it is understood that the letter is to be published in *Science* in spite of that journal's recent somewhat pompous strictures against authors seeking several outputs for a single manuscript.) What follows should therefore either be read as an anticipatory response to the argument that reductionists have no business bothering their heads with complicated questions such as the causation of cancer or should be regarded (in the British idiom) as an Aunt Sally, a gratuitous target of opportunity (in another idiom). The issue is nevertheless important, not merely because the problem of cancer is socially important but also because it bears directly on the design of an agenda for scientific enquiries of all kinds.

The public and sometimes even the

professional image of the reductionist is not flattering, a little like that of the out-and-out behaviourist among psychologists. A reductionist is one whose objective is to account for even the most complicated phenomena in terms of elementary interactions between more or less elementary entities. Elsewhere, among cosmologists for example, they may be people who worry about the phase changes that may account for the transition from a world in which free quarks exist to one in which all quarks are bound together in nucleons or sometimes mesons. In biology, reductionists appear as molecular biologists.

Everywhere, reductionists are intellectually vulnerable in three particular respects. First, and almost trivially, many obvious problems may lie technically beyond the reach of the reductionists: these days, it is a simple matter to calculate the electronic energy levels of a hydrogen atom, but simple molecules such as anthraquinone remain inaccessible except to those willing to include enough arbitrary constants in their calculations. Second, and more important, the reductionist agenda is no substitute for discovery: even in well trodden fields of physics such as the theory of wave motion, for example, it was not until recently appreciated that wave propagation can accommodate the motion of the apparently integral and unchanging packets of disturbance now called solitons. Is it any wonder that reductionist molecular biologists failed to predict reverse transcriptase, or the occurrence of introns in mammalian genes? Sensible reductionists, of course, acknowledge that their goal is not to reduce the whole of science to the definition of a few interparticle interactions and a mass of computation; the value of this way of looking at the world is rather to help more accurately to define the constantly moving boundary between what may be described as 'known' and the rest. Finally, reductionists are constantly exposed to the more subtle complaint that there are some phenomena so intricate, subtle and beyond the range of current concepts, that attempts to account for them serve only to demonstrate the arrogance, even foolishness, of those who try. Quite apart from the now discredited vitalists, reputable biologists belong in this third camp of critics. "Give us", they might say, "an atomic theory of memory and we might change our minds,"

The trouble with these arguments is that they are largely semantic. The difficulties of calculating the energy levels of anthraquinone do not undermine belief in the quantum theory but provide frequently convenient gibes, especially for those eager to suggest that some conceptual questions in science are forever inaccessible to the reductionists. But the motives by which those suggestions are inspired are themselves suspect. In relation to cancer, for example, it is often (and rightly) said that the most obvious characteristic of a cancer cell is that its specific differentiation has been to some extent unwound, so that it becomes an anarchic component of the tissue to which it belongs and also capable of finding a lodging in some quite different tissue. But then, the argument continues, molecular biologists have nothing tangible to say about differentiation (which is broadly speaking true). So how can they pretend that the time will soon arrive when a molecular explanation of malignancy is possible? Since, so far as can be told from our correspondent, the gibe stems not from a devotion to vitalism nor in a belief in other literally supernatural phenomena, the answer can be only another gibe — "Let's wait and see".

Balance

Meanwhile, it is important that the status of the reductionist agenda in relation to oncogenes should not be exaggerated. It would be folly to suggest that the time will ever come when the relationship between malignancy and molecular events is so well known and so predictable that oncologists will become people skilled only in the techniques of genetic manipulation. On the record of the past few centuries, but especially since Friedrich Wöhler found that ammonium cyanate will isomerize into urea, however, it seems risky to suppose that there are phenomena of any kind inaccessible to explanation in terms of simpler phenomena. And meanwhile, the practical issue between the reductionists and their critics is the all-important question of how the balance will be struck, in the understanding of cancer, especially as illuminated by what is being learned about oncogenes, between unexpected empirical discovery and the necessarily limited predictive power of the reductionists. The chances are that, as always, empirical discovery will make the running but that the causation of malignancy will be understood, possibly even dealt with. □

Child development

Phonological awareness: a preschool precursor of success in reading

from Max Coltheart

WORK reported by Lynette Bradley and Peter Bryant of Oxford University in this issue of *Nature* (p.419) provides the first clear evidence of the mental procedures important in the early stages of learning to read¹. A particularly striking part of their study shows that specific training methods can be applied before children have learnt to read which will considerably improve their later reading ability.

For some centuries it has been accepted that a child in the early stages of learning to read might be accomplishing the task of reading single words in either of two ways. One possibility is that the child uses a lexical or word-specific procedure which depends upon previously learned information about specific words — the printed letter string is used to access an entry in a developing mental lexicon, and a specification of the word's pronunciation is retrieved. A second possibility is that a non-lexical or rule-based procedure is used in which a learned system of rules relates particular letters or letter groups to particular spoken sounds — a system of rules converts a letter sequence to a sound sequence, which the reader then utters.

The work of Bradley and Bryant suggests that it is the non-lexical procedure which is the important one at first, that the child must attain competence in analysing spoken English words into their constituent sounds before beginning to learn to read using the non-lexical procedure, that children can be trained to attain such competence before they are old enough to be taught to read and that such training helps children to learn to read.

As lexical and non-lexical procedures are equally efficacious for the reading of the majority of English words, it is not possible to tell, merely from the fact that a child can read such words, which procedure is being used. Much research has been devoted to developing experimental techniques which make it possible to study the relative contributions of the two procedures.

One approach has been to compare young children who have made good progress in learning to read with those whose progress has been very slow. If the lexical procedure is important for young readers then a good reader should be adept at learning, in a one-to-one fashion, associations between particular abstract visual patterns and particular items of linguistic information (information about meaning or pronunciation). If, instead, the non-lexical procedure is more important for young readers, then a good reader would be expected to be someone with extensive knowledge of the rule relating letters to

their corresponding sounds. In an unpublished doctoral thesis, Firth² showed that good readers do indeed differ from bad readers in their knowledge of letter/sound rules rather than in their ability to learn linguistic associations to abstract visual patterns.

Reliance on the non-lexical procedure cannot, however, continue if the child is to become a skilled reader, for at least two reasons. The first is that the English language contains numerous words — the so-called 'exception' or 'irregular' words — whose pronunciation cannot be derived by using letter/sound rules, because the words do not obey these rules. The second is that the English language contains many homophones — words with different spellings but identical pronunciations. Since the conversion of letters to sounds makes homophonic words indistinguishable from each other, a reader relying solely on the non-lexical procedure could not determine whether 'sail' refers to ships or shops. The transition from early reading to skilled reading must therefore take the form of a transition from the use of the non-lexical procedure to the use of a lexical procedure which treats words as unique visual patterns. There is evidence that this transition occurs over the years 6 to 10 in normal readers³.

Given this view of learning to read, poor performance in reading, measured relative to other children of equal chronological age, could arise in either of two ways, depending on what the child's age is. A very young child who is poor at using the non-lexical procedure (letter/sound rules) will fare worse than his coevals in his initial attempts at acquiring the skill of reading, and it is precisely this difference that was shown by Firth². A child who is competent at the use of the non-lexical procedure, but finds it difficult to abandon the use of this procedure at a later age, will emerge as a bad reader at this later age and, specifically, will exhibit difficulties in reading irregular words and in discriminating between homophones. Precisely this pattern of poor reading in somewhat older children has been observed⁴.

If, for the first stage of learning to read, the learning of rules relating letters or letter groups to the individual sounds (phonemes) of English words is crucial, it will be necessary that the child be first capable of analysing spoken words into their constituent phonemes. A child who is poor at such phonological analysis should therefore experience difficulty in learning to read: and an association between poor ability to appreciate the phonological

structure of spoken English words and poor learning to read has frequently been reported in recent years⁵.

The existence of such an association is not sufficient, however, to show that poor phonological skills cause poor initial reading. To make such a claim would be to confuse correlation with causation. The association could well arise because in learning to read the child becomes aware of the phonological structure of English. It is also quite plausible that phonological ability and reading ability are not causally related at all, but that their association arises because, as Bradley and Bryant point out (see p.419), they are both related to some third variable (such as intelligence or social class).

Bradley and Bryant were, however, successful in tackling these problems. They were able to discount the first of the two alternatives by showing that children whose phonological skills were poor before the age at which reading instruction began made poorer progress in learning to read than children whose phonological skills were normal before learning to read.

The second alternative could not be discounted by this observation but was convincingly eliminated by a training study. A group of children who had not yet learnt to read was selected and randomly assigned to a training and a non-training condition. The first group was taught to become aware of the phonological structure of spoken English while the second was given instruction in another task — categorizing spoken words in terms of their meaning. The children went to school as normal and were exposed to conventional reading instruction. At the end of the training period reading ability was assessed and it was found that the training group learnt to read with considerably more success than the non-training group, showing that variations in the ease with which a child learns to read are caused by, rather than merely being associated with, variations in the ability to analyse phonological structure.

The work of Bradley and Bryant is thus important for two reasons: not only has it established a causal connection between phonological analysis and learning to read, but it also shows that there are methods for training phonological analysis which can be used with very young children and which have beneficial effects on the children's attempts to learn to read. □

Max Coltheart is Professor in the Department of Psychology, Birkbeck College, University of London, Malet Street, London WC1E 7HX.

1. Bradley, L. & Bryant, P.E. *Nature* **301**, 419 (1983).
2. Firth, I. *Components of Reading Disability* thesis, Univ. New South Wales (1972).
3. Doctor, E.A. & Coltheart, M. *Memory Cognition* **8**, 195 (1980).
4. Coltheart, M., Masterson, J., Byng, S., Riddoch, J. & Prior, M. *Q. Jl exp. Psychol.* (in the press).
5. Jorm, A.F. *The Psychology of Reading and Spelling Difficulties*, Ch.4 (Routledge & Kegan Paul, 1983).

Migratory DNA

Mitochondrial genes in the nucleus

from Thomas D. Fox

No matter how one thinks about the evolution of mitochondria and their intra-organellar genetic systems, one is invariably forced to suppose that movement of genes occurs, or has occurred, between the mitochondrion and the nucleus. Several studies are now revealing that the nuclei of yeast, *Podospora*, sea urchins and locusts contain sequences homologous to clearly defined mitochondrial genes and further substantiate the notion that mitochondrial DNA (mtDNA) can travel within the cell. Further, it appears from the degree of homology observed in apparently functionless pieces of DNA that such transfers of mtDNA have occurred relatively recently in the evolutionary past.

One of the clearest demonstrations of mtDNA pseudogenes in the nucleus comes from work on the yeast *Saccharomyces cerevisiae* that has just appeared in *Nature*. Farrelly and Butow¹ used a cloned piece of mtDNA that encodes the mitochondrially synthesized ribosomal protein var1 to probe Southern blots of nuclear yeast DNA isolated from q^- strains that contain no *bona fide* mtDNA. They found that the probe hybridized specifically to a single restriction fragment in all strains examined, although the strains fell into two classes based on the size of the hybridizing fragment. The fact that a sequence homologous to mtDNA is present in a strain lacking mtDNA does not necessarily prove that that sequence is located on a nuclear chromosome. Farrelly and Butow took advantage, however, of the size polymorphism by crossing strains from the two size classes and showing that the polymorphic fragments segregated in the cross as mendelian traits, directly showing their chromosomal location.

Analysis of the cloned nuclear DNA fragments by hybridization and sequencing revealed a number of interesting features. First, the size polymorphism is due to the presence, in some strains, of two yeast transposable elements (Ty) oriented tandemly close to the mitochondrial sequence. Second, the nuclear DNA is homologous not only to the *var1* gene but also to part of the *cob* gene (containing both intron and exon sequences) and a region recognizable as a putative mitochondrial origin of replication. These sequences are not contiguous in mtDNA, and their pattern of rearrangement resembles that found in q^- mutants of yeast, which retain amplified and often rearranged portions of the mitochondrial chromosome. Finally, the actual degree of sequence homology between the nuclear DNA and the related mtDNA sequences is quite high (80 per cent). Since no plausible

function can be assigned to these nuclear sequences (based on the fact that they are only partial genes with multiple frameshifts and in any event are written in yeast mitochondrial genetic code) it can be reasonably assumed that they have diverged freely since their incorporation into the nucleus. Based on the observed sequence divergence Farrelly and Butow estimate that the nuclear pseudogene diverged from the present mitochondrial gene about 25 Myr ago; rather recently when viewed against the evolutionary history of modern fungi, not to mention eukaryotes in general.

Another fungus in which rearrangement and amplification of mtDNA sequences occurs is *Podospora anserina*, a filamentous ascomycete that undergoes mycelial senescence as part of its life cycle. The amplification of particular mtDNA sequences (termed senDNA) is associated with and apparently causally related to the phenomenon of senescence. A report from Wright and Cummings², shortly to appear in *Nature*, shows that hybridization of cloned senDNA to Southern blots of nuclear DNA purified by density gradient centrifugation from senescent mycelia, reveals nuclear sequences homologous to the mitochondrial senDNA. Although no *Podospora* strains totally lack mtDNA, a mutant that escapes senescence, as a result of a mitochondrial deletion removing a senDNA sequence, has been isolated³. Wright and Cummings report that even in the mutant lacking this senDNA sequence in the mitochondria, homologous sequences are present in the nuclear DNA. They speculate that the movement of senDNA sequences from the mitochondria to the nucleus may be a normal part of the senescence process, although no direct evidence for this is presented. Cloning the nuclear DNA sequences homologous to mtDNA will be required to analyse the system further.

Mitochondrial DNA sequences are present not only in the nuclei of fungi, but have also been discovered in the chromosomal DNA of sea urchins and locusts. In both cases the sequences were detected by probing whole-genome Southern and genomic libraries with cloned cDNA corresponding to mitochondrial transcripts. In the case of the sea urchin (*S. purpuratus*), Jacobs *et al.*⁴ describe a cloned chromosomal fragment that contains sequences homologous to both the cytochrome oxidase subunit I gene and to a portion of the mitochondrial 16S rRNA as judged by hybridization and partial sequence analysis. (Other regions of homology were detected but not studied further.) The

cytochrome oxidase gene is in fact present twice; one possibly complete copy and a second truncated version. Between them (and contiguous with the full-length oxidase gene) lies a small portion of the rRNA sequence. Jacobs *et al.* suggest this is the result of an integration and a partial duplication of the mtDNA sequence, although a yeast scientist can hardly help wondering whether this is not a kind of animal q^- event preserved in the nucleus (no such rearrangements have ever been detected in animal mitochondria). Melting studies on heteroduplexes formed between mitochondrial sequences and the homologous nuclear DNA yielded the surprising result that the sequence divergence of the 16S rRNA sequences (8 per cent) was rather less than the cytochrome oxidase subunit I sequences (13–25 per cent depending on the region). The reason for this is unclear. The authors suggest that perhaps some form of selective pressure is affecting parts of this region. If it is assumed, however, that at least some of these sequences can mutate freely, the divergence time can be estimated to be 25 Myr or less. In the case of the locust (*Locusta migratoria*), Gellissen *et al.*⁵ isolated several distinct clones of nuclear DNA, each of which has detectable homology to mtDNA and to mitochondrial rRNA. While the detailed arrangement of these homologous sequences remains to be worked out, the results so far show that in locusts, as in sea urchins, at least some mtDNA sequences are not only present but indeed repeated in the nuclear genome.

Another instance of movement of a gene, that coding the ATPase subunit 9, from the mitochondria to the nucleus or vice versa during the evolution of *Ascomycetes* has been known for some time. In yeast the active gene for this protein is in the mitochondria⁶. The *Neurospora* gene is nuclear⁶, although an apparently silent gene encoding a homologous protein is also present in *Neurospora* mitochondria⁷. Are the mechanisms that move an active gene between an organelle and the nucleus similar to those that transport mitochondrial sequences to apparent retirement in the nucleus? It is easy to imagine that DNA (or RNA) released into the cytoplasm during breakdown of a mitochondrion would find its way into the nucleus, as in standard transformation procedures. Movement in the other direction might be less 'easily' achieved,

1. Farrelly, F. & Butow, R.A. *Nature* **301**, 296 (1983).
2. Wright, R.M. & Cummings, D.J. *Nature* (in the press).
3. Vierny, C., Keller, A.-M., Begel, O. & Belcour, L. *Nature* **297**, 157 (1982).
4. Jacobs, H.T. *et al.* *J. molec. Biol.* (in the press).
5. Gellissen, G., Bradfield, J.Y., White, B.N. & Wyatt, G.R. *Nature* (in the press).
6. Sebald, W. & Wachter, E. in *29th Mosbacher Colloq. on Energy Conservation in Biological Membranes* (eds Shaefer, G. & Klingenberg, M.) 228 (Springer, Berlin, 1978).
7. van den Boogaart, P., Samallo, J. & Agsteribbe, E. *Nature* **298**, 187 (1982).
8. Stern, D.B. & Lonsdale, D.M. *Nature* **299**, 698 (1982).

although the recent demonstration that a substantial piece of chloroplast DNA is present in the mtDNA of maize⁸ shows that DNA can get into mitochondria. Real experimental analysis of these evolutionarily important processes will

have to await the development of techniques for the genetic transformation of organelle genomes. □

Thomas D. Fox is at the Section of Genetics and Development, Cornell University, Ithaca, New York 14853.

X-ray astronomy

Hot astrophysical plasmas

from Richard McCray

BEGINNING with the launch of the UHURU satellite in 1969, more than 10 X-ray detectors have been flown on satellites to create a revolution that has affected virtually all areas of astronomy. High-temperature ($> 10^6$ K) plasmas have been observed in most types of stars, interstellar and intergalactic gas, compact binary sources, active galactic nuclei and distant quasars. The outstanding scientific questions of X-ray astronomy were reviewed at a recent conference* held in Nice, with the aim of developing a strategy for future European space missions.

At present, X-ray astronomy is experiencing a hiatus with only one relatively modest instrument, the Japanese *Hakuchō* satellite, in orbit. The situation will improve in summer 1983 with the launches of the EXOSAT mission by the European Space Agency (ESA) and of the ASTRO-B mission by Japan. Beyond this, the only X-ray satellites under construction are the Japanese ASTRO-C (launch expected in 1986) and the German ROSAT (launch expected in 1987).

Studies of laboratory plasmas complement those of astronomers, as was evident in the review (R. Mewe, Space Research Laboratory, Utrecht) of X-ray emission by optically thin plasmas, including interstellar matter, supernova remnants, solar flares and Tokamak plasmas. In each case the X-ray spectra are rich in emission lines, the interpretation of which may yield the temperature, element abundances and density of the gas. The main thrust of current theoretical research is to understand how the X-ray emission spectrum is affected by the departures from the static ionization balance likely to occur in these environments.

The X-ray spectra of supernova remnants (H. Itoh, Institute of Astronomy, Cambridge) show, surprisingly, that the abundance of Si and S are enhanced much more than that of iron (which was expected to be the main product of nucleosynthesis). Furthermore, supernova shells do not seem to decelerate according to the Sedov blast wave model that has long been a mainstay for their interpretation. Perhaps the interstellar medium around the supernova has been removed by the pre-supernova star or previous supernovae.

X-ray emission from intergalactic gas in clusters of galaxies (A. Cavaliere, Institute

of Physics, Rome; R. Mushotzky, NASA-Goddard Space Flight Center) contains emission lines of Fe, Si and S which imply that the abundances of heavy elements are comparable with those in the solar neighbourhood, indicating that the gas is not primordial but comes from the galaxies themselves. Furthermore, a few rich clusters show evidence for cooling flows towards the dominant central galaxies at a rate of a few hundred solar masses per year. The major challenges now are the processes that remove gas from the galaxies and control the flow of the intergalactic gas, and the consequences for galactic evolution.

X rays have been observed from almost every type of star (except red giants), often with luminosities far exceeding that of the solar corona. In the case of hot supergiant stars (C. De Loore, University of Brussels), the X-ray emission seems to indicate hydrodynamic instabilities in the powerful winds of these stars. In the case of low mass-stars (J. Heise, Space Research Laboratory, Utrecht), the X-ray emission suggests magnetic activity similar to that on the solar surface, but often on a much greater scale. Intense and rapidly varying coronal activity seems to be a characteristic of young stars and may provide important clues to star formation.

Compact galactic X-ray sources, including accreting white dwarfs, neutron stars and black holes in binary systems, are characterized by rapid X-ray variability and complex spectra. Galactic compact binary sources (S. Ilovaisky, Observatoire de Besançon; J. van Paradijs, Astronomical Institute, University of Amsterdam) fall into two distinct groups. Population I sources have early-type massive companion stars and hard X-ray spectra; most are X-ray pulsars and almost certainly are magnetized neutron stars. The enigmatic population II sources (including X-ray bursters) have low-mass companions, soft X-ray spectra and do not pulse; most are probably neutron stars with weak magnetic fields.

Quasars and active galactic nuclei (A.C. Fabian, Institute of Astronomy, Cambridge) have hard X-ray spectra extending in some cases to > 100 keV. The rapid variation in their enormous X-ray luminosity approaches and even appears to exceed fundamental theoretical limits. As with the 'superluminal' motions of radio lobes, however, the problem may be

resolved if the X-ray-emitting regions have relativistic bulk motions. The presence (in Seyferts and quasars) or absence (in BL Lac objects) of broad emission line regions may be related to the spectral index of the central X-ray source, as a consequence of the equation of state of gas illuminated by X rays. X-ray spectra of some active galactic nuclei (L. Culhane, Mullard Laboratory) indicate photoabsorption by gas clouds near the nucleus.

The scientific discussions all pointed to a need for X-ray telescopes with substantially increased collecting area and capability for spectroscopy. To understand supernova remnants and clusters of galaxies we need the capability to obtain two-dimensional images of the sources with moderate ($\lambda/\Delta\lambda \sim 100$) resolving power. To understand stellar coronae we need higher ($\sim 1,000$) spectral resolution. Observations of time variability of the spectra of compact binary sources are needed to unravel the complex problems of spectral formation, radiative transfer and gas dynamics. Higher sensitivity will provide a greater sample of active galactic nuclei and better statistics on the time variability of their spectra, both in soft (< 2 keV) X rays where absorption by gas clouds is evident and in hard (~ 100 keV) X rays where most luminosity appears.

Future plans for X-ray astronomy are still unsettled. The technology for constructing X-ray mirrors, spectrometers, and detectors has advanced to a stage where it is feasible to build X-ray telescopes with vastly improved sensitivity and spectroscopic capability. No nation has made, however, a commitment to build them. The US AXAF telescope will not be launched until the early 1990s. The 1987 ROSAT mission will conduct a sensitive all-sky survey of more than 10^5 sources, but lacks spectroscopic capability.

There is no shortage of exciting plans. British astronomers have proposed a mission with a resolving power of ~ 50 –500 from 100 eV to 10 keV, while an Italian proposal is for a telescope with relatively low spectral resolution but extended (2–200 keV) spectral range. The ESA CORONA mission could map the diffuse soft X-ray background with moderate (~ 100) spectral resolution. The versatile X-80 mission proposed for ESA could address virtually all the scientific problems discussed at the meeting. For the mid 1990s, European astronomers have proposed a spectroscopic mission, MMX, of truly awesome size and power. In 1983, ESA and Japan will take over the leadership in X-ray astronomy from NASA. By building powerful spectroscopic missions such as those proposed at the conference, European astronomers could maintain this leadership for a long time. □

Richard McCray is Professor of Astrophysics at the Joint Institute for Laboratory Astrophysics of the University of Colorado and the National Bureau of Standards, Boulder, CO 80309.

*The meeting was held at Nice, France on 8–10 November 1982. Proceedings will be published in *Physica Scripta*.

El Chichón eruption

The dust cloud of the century

from Alan Robock

THE 4 April 1982 eruption of the El Chichón volcano in Chiapas state, Mexico, produced a stratospheric dust veil that is higher and probably more massive than any this century. Local temperature increases in the stratosphere of more than 5°C have already occurred and a cooling of about 0.5 °C is expected at the Earth's surface. The high sulphur content of the volcanic gases, and subsequent conversion to small, bright, long-lasting sulphuric acid particles, ensures that the effects will be large and continue for several years.

The vast array of observations being taken from the ground, aeroplanes, balloons and satellites (see the figure) will ensure that it is the best observed and monitored volcanic cloud ever. The eruption has provided meteorologists with an unprecedented opportunity to test theories of gas-to-particle conversion, stratospheric transport and settling of dust particles, radiative interactions of dust particles and resulting changes of stratospheric and surface air temperatures. At two recent meetings* scientists discussed the continuing evolution of the dust cloud and its climatic significance. Attention focused on the following topics: the three-dimensional location of the cloud, the changing gaseous and particulate composition of the cloud, observations of possible weather and climate effects, and theoretical models of the above processes.

Using images from the GOES and NOAA-7 weather satellites, M. Matson (NOAA) described the sequence of the eruptions. The first occurred on 29 March (GMT) and probably produced a small stratospheric cloud at a height of about 20 km. Measurements of the thermal IR emissions from the cloud allowed estimates to be made of temperature and comparisons with nearby radiosonde measurements of the atmospheric temperature profile then allowed estimates of height. Comparisons of the direction of motion of the cloud with winds at different levels confirmed a stratospheric component moving towards the southwest. Two smaller non-stratospheric eruptions on 3 April were followed by the large eruption on 4 April which produced the massive stratospheric cloud later shown to be densest at about 26 km. A. Krueger (NASA) also observed the initial eruption sequence, using the Total Ozone Mapping Spectrometer (TOMS) on the Nimbus-7 satellite. Anomalous ozone measurements turned out to be the UV signature of sulphur dioxide gas in the eruption cloud and allowed him to track and measure the total sulphur dioxide.

Matson and A. Robock (University of Maryland) used the weather satellites to

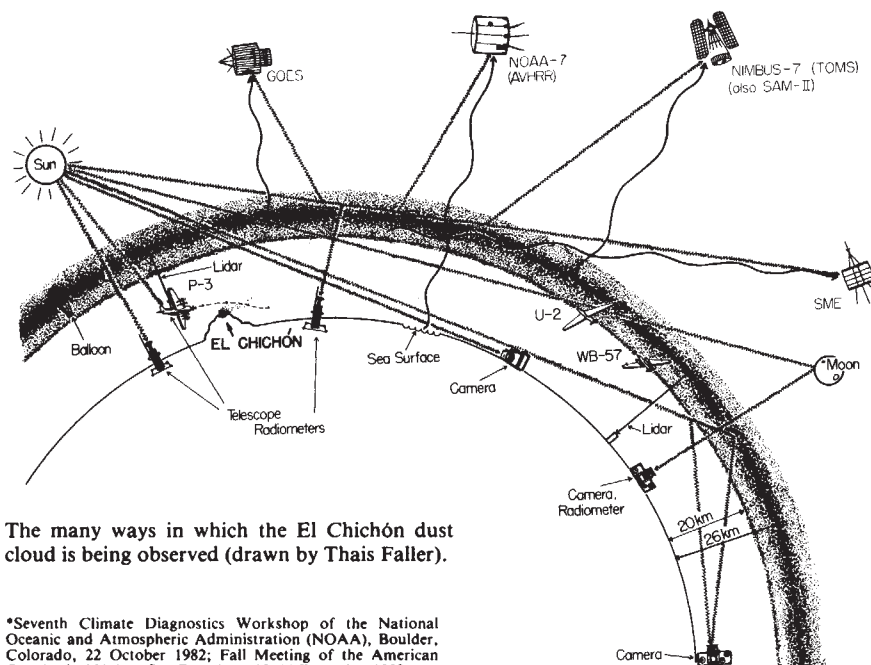
track the visible image of the dust cloud as it travelled due westwards, completely circling the globe in 21 days at a mean speed of 22 m s⁻¹. That it could be seen for this long on satellite photographs indicates just how massive the cloud was. The Mount St Helens cloud, in contrast, could only be followed for 3 or 4 days after the eruption. R. West (University of Colorado) found identical results from observations of scattered sunlight taken with the limb-scanning visible spectrometer on the Solar Mesosphere Explorer (SME) satellite.

Fortunately, the densest part of the 4 April cloud went right over Hawaii, arriving on 9 April, allowing many different surface measurements of the cloud. Hawaii is one of a few places in the world with vertically pointing lidar facilities. Lidar (acronym for Light Detection And Ranging) works similarly to radar, but sends up a laser light pulse and uses a telescope to measure the timing and intensity of the light scattered back from particles in the atmosphere. K. Coulson (Mauna Loa Observatory) reported that the Hawaii lidar measured the highest and largest cloud ever seen, with backscatter ratios (total to gaseous) of over 200. The 'mystery cloud' which came from an as yet unidentified eruption early in 1982 had, by contrast, backscatter ratios of about 3 at the most. The lidar saw finely structured layers of particles with a small peak at about 20 km and the large peak at 26 km, but with some dust over 30 km. Later measurements showed reduced backscatter ratios at the peak, but larger values above the peak and less structure to the layers,

indicating continuing particle formation and horizontal diffusion. Other lidar measurements, made around the world, allowed the latitudinal spread and total amount of the cloud to be monitored over time.

It also became possible to locate the El Chichón cloud because it blocked some of the outgoing long-wave radiation being measured in a new programme designed to investigate global sea-surface temperatures (A. Strong, NOAA). The differences between the satellite sea-surface temperature measurements and ship and buoy measurements at the surface gave maps of the volcanic cloud's location and relative thickness. Other measurements from the SME satellite in longer-wavelength bands (G. Thomas and R. Sanders, University of Colorado) also showed the vertical and horizontal extent of the cloud. In addition, McCormick reported measurements of the cloud over the polar regions with the SAM II instrument on the Nimbus-7 satellite, and R. Keen (University of Colorado) used observations of lunar brightness during an eclipse to infer the latitudinal distribution of the dust. *In situ* balloon measurements of particle density in the cloud, discussed by D. Hofmann (University of Wyoming), also located the cloud.

Except for the SME results in visible wavelengths, all the above measurements showed that the upper cloud quickly spread, occupying the latitude band between the equator and 30°N after 2 months, and then spread very slowly to the north and the south so that it now occupies a band from about 10°S to 35°N. The lower 20 km cloud has probably spread to completely cover the globe. The SME visible results showed particles to above 40 km in the Northern Hemisphere and large concentrations in the Southern



The many ways in which the El Chichón dust cloud is being observed (drawn by Thais Faller).

*Seventh Climate Diagnostics Workshop of the National Oceanic and Atmospheric Administration (NOAA), Boulder, Colorado, 22 October 1982; Fall Meeting of the American Geophysical Union, San Francisco, 10–11 December 1982.

Hemisphere that were not observed by any other means. The reasons for these discrepancies are not yet understood.

It has become accepted, as pointed out by S. Self (Arizona State University) and B. Toon (NASA), that sulphur gas content is the important factor determining the long-term climatic effects of eruptions that place material in the stratosphere. Silicate particles are large, fall out rapidly and have only a short-term effect. Sulphur gases, on the other hand, are converted to sulphuric acid particles which are very small and thus have a long mean residence time in the stratosphere — around 2 or 3 years. The particles are very bright and have a large effect on the amount of solar radiation received at the Earth's surface by scattering several per cent of the radiation back to space. It is therefore important to measure the silicate to sulphur ratio of the particles and the concentrations of the sulphur gases that are precursors to the particles, as functions of time and space.

Several NASA research aeroplane flights reached heights of about 20 km and were able to sample the lower cloud, and samples were also taken of gases currently being emitted by the volcano. Only balloon flights were able to take samples from the main 26 km cloud. High-altitude aeroplane collections have been examined by scanning electron microscopy and show silicate particles with distinct sulphuric acid pitting. Energy-dispersive X-ray spectrometer analysis of the particles showed sulphur was a significant component. Particles collected by balloon from within the main cloud were at least 95 per cent sulphuric acid. Krueger measured only 3.3 million tons of sulphur dioxide with the TOMS instrument in the eruption cloud on 6 April, but calculations of the total mass based on later observations of the dust particles suggest that more sulphur gas should have been put in by the volcano. Carbon sulphide was not found by the high aeroplanes, but recent samples made at the volcano showed emissions rich in hydrogen sulphide. If the actual eruption cloud had the same composition, then it would take some time for the hydrogen sulphide to convert to sulphur dioxide and then to sulphuric acid particles.

Radiative effects of the dust cloud were seen in ground measurements of incoming solar radiation and optical effects in the sky as well as satellite measurements of the Earth's radiation budget. B. Mendonca (NOAA) and E. Yasukawa (Mauna Loa Solar Observatory) reported that the direct solar radiation at Hawaii was reduced by more than 20 per cent and the total by more than 7 per cent. These are the largest reductions of direct solar radiation ever observed there. J. DeLuise (NOAA), W. Lockwood (Lowell Observatory) and B. Herman (University of Arizona) all used radiation measurements to estimate the particle size distributions of the cloud. DeLuise calculated optical depths of the cloud and estimated that its total was about

Immunological diversity

How the elephant got its trunk

from Elizabeth Simpson and Michael Pope

A PAPER in this issue of *Nature*¹ (p.388) presents some of the first evidence that gene conversion or homologous recombination may be used to generate a part of the striking polymorphism of transplantation antigens.

Multicellular organisms must distinguish between self and non-self in order to deal with viruses and other harmful foreign invaders. In mammals the immune system has evolved complex defence mechanisms which centre round the activities of B and T lymphocytes. B lymphocytes produce antibodies which bind specifically to foreign determinants (antigens) whereas T lymphocytes recognize antigens only when they are presented together with self, class I or II molecules of the major histocompatibility complex (MHC). Almost all cells of the body have on their surfaces class I MHC molecules; these tissue antigens are also the most important in triggering graft rejection when tissues of one individual are transplanted into a non-identical individual of the same species.

A striking feature of class I MHC molecules is their extraordinary polymorphism — in the mouse there are at least 100 alleles at the principal class I loci. Such polymorphism almost certainly has selective advantage for the species. It has been a matter of speculation, though, as to how this polymorphism arises at the gene level. One hypothesis was that each individual possessed copies of genes for all possible 'alleles' found in the species but that expression was selective. This possibility has been rendered unlikely by the finding that there are too few genes at the principal class I loci (*H-2K*, *D* and *L* in the mouse).

15 million tons.

The radiative effects of the dust cloud and resulting surface air temperature changes were modelled by R. Turco (R&D Associates), J. Pollack (NASA), F. Luther (Lawrence Livermore National Laboratory) and Robock, each using different types of climate models. All showed cooling, in fair agreement with each other. Robock's model gave time-dependent latitude-dependent results, predicting Northern Hemisphere average cooling of about 0.5 °C during 1984 and 1985 with the effect becoming less after that. Local stratospheric warming of more than 5 °C has already been observed (J. Angell and R. Quiroz, NOAA; McCormick). It is difficult, however, to separate the signal from the normal quasi-biennial temperature variations in the stratosphere that are of about this amplitude. Strong made a controversial

Another possibility is that of gene conversion, by which exons of adjacent homologous genes are recruited into the expressed gene. Analogies exist in bacterial², lower eukaryotic³ and viral systems⁴.

Evidence consistent with such a mechanism comes from recent findings of Mellor *et al.*⁵ who have cloned and sequenced *H-2K^b* and mutant *H-2K^{bm-1}* molecules and found fairly extensive DNA sequence alteration in the mutant sequence, suggestive of gene conversion rather than point mutation at the DNA level. The work reported in this issue¹ shows that isolated partial class I DNA sequences can be inserted and expressed by cellular DNA. DNA clones containing only the first three exons of either a *K^d* or *L^d* gene (the complete gene includes eight exons), were inserted into L cells which then expressed complete *K^d* or *L^d* molecules on their surfaces, presumably by utilizing their own homologous exons 4–8.

Perhaps class I MHC genes commonly use such gene rearrangements to guarantee their remarkable genetic diversity and possibly this reflects a general strategy by which other polymorphic gene clusters may produce their variability. □

Elizabeth Simpson is in the Transplantation Biology Section and Michael Pope in the Clinical Sciences Division of the Clinical Research Centre, Watford Road, Harrow, Middlesex HA1 3UJ.

1. Goodenow, R. S. *et al.* *Nature* **301**, (1983).
2. Radding, C. M. *Rev. Biochem.* **47**, 847 (1978).
3. Stahl, F. W. *Genetic Recombination* (Freeman, San Francisco, 1979).
4. Winocour, E. & Keshet, I. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929 (1980).
5. Varmus, H. E., Quintrell, N. & Ortiz, S. *Cell* **25**, 23 (1981).
6. Mellor, A. *et al.* *Nature* (in the press).

suggestion that the warming was responsible for a weakening of the Hadley cell circulation, resulting in anomalously warm equatorial Pacific sea-surface temperatures and hence affecting the winter climate of North America. His idea is intriguing, but remains to be proved. A model proposed by L. Capone (San José State University) suggested that large anomalies in the stratospheric circulation induced by the warming would explain the high altitude of the cloud as well as its slow spread northwards. El Chichon in Spanish means 'lump on the head', and I am sure the modellers, myself included, will be prepared to accept our chichones should our results prove faulty. □

Alan Robock is an Associate Professor in the Department of Meteorology, University of Maryland, College Park, Maryland 20742.

X-chromosome inactivation

Switching off blocks of genes

from Lucio Luzzatto and Stanley M. Gartler

THE inactivation of one of the two X-chromosomes in the somatic cells of female mammals is a developmental event that dramatically affects the expression of a large block of genes. How, at the molecular level, inactivation is brought about is a riddle that has so far defied complete solution.

Three alternative kinds of change could account for the difference between an active and an inactive X-chromosome: DNA sequence changes, DNA modification and changes in the chromatin not involving DNA. Although there is evidence for and against each of these alternatives, none has yet been formally ruled out. At a recent meeting* on the topic it appeared, however, that several lines of evidence are converging to indicate DNA modification, probably involving methylation, as the most likely alternative. In addition, data on the structure of three well identified human X-chromosome genes were presented for the first time at the meeting.

The strongest evidence that X-chromosome inactivation involves a change in DNA is the finding that DNA taken from inactive X-chromosomes will not transform mammalian cells deficient in hypoxanthine phosphoribosyl transferase (HPRT) production — an enzyme whose activity is readily detected and which is coded for on the X-chromosome — while DNA taken from the active X-chromosome will transform the same cells (Liskay, R.M. & Evans, R.J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4895; 1982).

This work has now been confirmed in a much more extensive series of experiments (R.M. Liskay, Yale University), although A. Westerfeld (Erasmus University) presented contrary evidence in a human-mouse system (*Nature* **295**, 624; 1982). At the same time, S.M. Gartler (Washington University) reported new data indicating that DNA from cells from a woman heterozygous for HPRT deficiency could only function in transformation of HPRT-mouse cells when the normal allele was on the active X-chromosome. The conflict in the experimental results was not resolved.

On the other hand, Liskay also reported that inactive X DNA from extra-embryonic trophoblast tissue does function in HPRT transformation. It thus seems likely that X-inactivation in this case differs from that in the embryo itself not only in being nonrandom (the paternal X is always inactivated, whereas in the embryo either the paternal or the maternal

X-chromosome is inactivated at random), but also in its molecular mechanism. It is possible that this intriguing development may be explored further in the nonrandom pattern of X-inactivation described by G. Martin (University of California, San Francisco) in murine teratocarcinoma, as well as in marsupials.

Interestingly, the three human X-specific genes for which cloned sequences have been obtained — glucose 6-phosphate dehydrogenase (G6PD), HPRT and phosphoglycerate kinase (PGK) — all map in the subtelomeric region of the long arm of the chromosome. (A fourth gene, specifying coagulation factor IX, the location of which on the X-chromosome is unknown, has also been recently cloned by G. Brownlee's group in Oxford: *Nature* **299**, 178; 1982). G. Persico (IIGB, Naples) reported on a 600-base pair long clone from fibroblast G6PD cDNA (*Nature* **294**, 778; 1981) which has now been sequenced and shown to correspond to the 3' end of the mRNA. It does not cross-hybridize in stringent conditions with either mouse or hamster DNA, but it does hybridize with DNA from a human-mouse cell hybrid containing only the Xq26 $\rightarrow$ ter portion of the human X. By sequencing, it was found that the mRNA contains at least 1,000 bases beyond the stop codon and at least one intron near the stop codon.

Regarding HPRT, T. Caskey (Baylor College of Medicine, Houston) reported

the cloning of the gene after it was 'amplified' in aminopterin-resistant cells. For a polypeptide chain of about 26 K the gene turns out to be very large, with an estimated span of 33 kilobases containing five introns (see also Jolly, D.J. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 5038; 1982). One of the cDNA clones obtained has been sequenced. An interesting feature of this probe is that although it undoubtedly recognizes an X-linked gene, it also cross-hybridizes with autosomal sequences (are there translocated HPRT 'pseudogenes'?). At least one *Bam*HI polymorphism has been detected by testing a few dozen individuals, and in one patient with HPRT deficiency (causing the severe neurological disorder called Lesh-Nyhan syndrome) an abnormal *Bam*HI pattern was observed, suggesting a possible deletion in the gene.

The third gene, PGK, was cloned by synthesizing a set of oligonucleotides corresponding to six amino acids of the known protein sequence, and using them to screen a cDNA library from an adenocarcinoma cell line (J. Singer and A. Riggs, City of Hope Research Institute, Duarte, California).

One classical feature of X-inactivation, other than in marsupials, was, with a single exception (*Proc. natn. Acad. Sci. U.S.A.* **72**, 1510; 1975), its irreversibility. This taboo has now been abundantly desecrated with the report of a spontaneous G6PD 'reactant' in a diploid human fibroblast clone (B. Migeon, Johns Hopkins University) and the reactivation of an inactive HPRT gene induced by 5-azacytidine (5-azaC) (R. DeMars, University of Wisconsin, Madison and earlier reports from Shapiro's group: *Science* **211**, 393; 1981). Because 5-azaC

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

* A workshop on 'Molecular Biology of X-chromosome Expression and Inactivation' was organized by S. Gartler of the University of Washington, Seattle, and by F. Biali and D. Toniolo of the International Institute of Genetics and Biophysics, Naples; it was sponsored by the US National Science Foundation and by the Consiglio Nazionale delle Ricerche of Italy, and was held in Amalfi, 1-3 November 1982.

THE above photograph of a high-voltage electrical discharge, taken in 1892 with a lens-less camera, represents one of the many achievements of AA Campbell Swinton (1863-1930), the originator of the all-electronic concept for the realization of television. Campbell Swinton's career — and his pioneering work with X rays, wireless and television — is the subject of a new monograph by T.H. Bridgewater (Royal Television Society monogr. 1, 2nd edn. The Royal Television Society, 1982).

affects DNA methylation and because 5-azaC-induced reactants vary in a permanent way in their level of HPRT expression, this result was interpreted in terms of the existence of multiple methylation sites which, depending on the degree of methylation, can exert quantitative control over the level of expression.

On the short arm of the human X-chromosome, two genes are conspicuous because they seem to escape inactivation: the blood group gene *Xg^a* and the steroid sulphatase gene (STS). That the STS locus on the inactive X is only partially expressed, rather than being fully shut down, was shown by Migeon in cloned human fibroblasts from a heterozygote for STS deficiency (see *Nature* 299, 839; 1982) and by T. Mohandas (University of California) in a mouse-human cell hybrid in which an inactive structurally abnormal human X-chromosome was retained and yet human STS activity was expressed.

In the end, it is possible that two distinct molecular events will emerge as responsible for X-inactivation in mammals. The cytological evidence and the data on spreading of inactivation point to an initial cooperative phenomenon involving most or all of the chromosome. But the evidence from spontaneous or artificial reactants suggest that the maintenance of inactivation is a local phenomenon associated with individual genes. This is reminiscent of dosage compensation in *Drosophila*, and it is tempting to think of maintenance of hyperactivity in males or hypoactivity in females as an ancient mechanism in evolution, on which block inactivation was subsequently superimposed as an added phenomenon in mammals. □

Lucio Luzzatto is in the Department of Haematology, Royal Postgraduate Medical School, University of London, London W12 0HS and Stanley M. Gartler in the Department of Genetics, University of Washington, Seattle.

Cystic fibrosis

One part of the puzzle

from V. Schwarz

In a paper published in this issue of *Nature* (p.421), Paul Quinton of the University College of Los Angeles Medical School reports measurements of potential differences across the isolated sweat duct epithelium of patients suffering from cystic fibrosis and shows that the tissue has an unusually low permeability to chloride ions compared with normal sweat duct. The failure to reabsorb sodium chloride and the high concentration of salt in the patients' sweat — a characteristic of patients with cystic fibrosis — may thus be explained.

The main symptoms of cystic fibrosis suggest that it is due to a malfunction of electrolyte and water transport in exocrine glands. The mucous glands are most seriously affected, while the sweat glands, which produce little mucus, are free from any discernible pathology; yet the sweat electrolyte defect is one of the earliest and most common manifestations of the illness, being demonstrable soon after birth, often before the appearance of the clinically more serious secondary features.

So far the primary genetic defect has eluded researchers, although much effort has been expended in the past thirty years. Researchers have not been at a loss to put forward hypotheses, frequently based on unexplained or unrepeatable observations, and per kilo of printed matter, cystic fibrosis research must surely embody more disagreement among investigators than most other subjects (see, for example, P.J. Benke in *Cystic Fibrosis*, eds A. Mangos & R. C. Talamo, 1977).

In the absence of a satisfactory animal model, attention has focused in turn on patients' urine, faeces, serum, sweat, saliva, semen, breast milk, blood cells,

skin, finger nails, nasal mucosal explants, rectal biopsies and cultured fibroblasts, and all have yielded publishable findings. The reported abnormalities are now so numerous and so diverse as to defy invention of a genetic 'umbrella' under which they can be accommodated.

In this chaotic situation some researchers have eschewed the siren call of easily cultured fibroblasts and readily available tissue fluids and have clung steadfastly to a tissue which is both simple in its anatomy and straightforward in its physiology, as well as grossly abnormal in its function, yet free from secondary pathology — sweat glands.

The normal sweat gland produces water for evaporation by first pumping NaCl into the lumen of the upper secretory portion of the gland; the salt is followed by water to form an isotonic precursor fluid which passes along the proximal duct (coiled around the secretory part) where most of the salt is actively reabsorbed from the lumen, leaving a hypotonic sweat to reach the surface of the skin. The beauty of the arrangement is that NaCl, used as a 'water carrier', cycles between secretory coil and duct, and only about 10–20 per cent is lost in the sweat. In cystic fibrosis the precursor fluid is normal but ductal reabsorption is defective and thus salt concentrations of 70–140 mM are common in the surface sweat, compared with a normal concentration of 15–25 mM. The loss of salt is clinically of little consequence unless the patient is exposed to excessive heat (it was discovered in a New York heat wave), but the defect is an invaluable diagnostic sign.

The seemingly simple process of salt

reabsorption in the normal sweat duct is so far ill understood and several questions remain to be answered or even asked. For instance, does the diffusion of Na⁺ and Cl⁻ into the duct cell, before active transport into the extracellular fluid, limit the overall rate of salt transport? If so, is there some mechanism for increasing the basal permeability when the duct is required to reabsorb salt from the precursor fluid? Might such regulation be under neurohumoral control? Are there any surface receptors to transduce signals received from other parts of the body or even of the gland? The duct is known to respond to aldosterone in the same manner as the renal tubule, albeit more slowly — a facility unimpaired in cystic fibrosis; but the possible intervention of other hormones remains unexplored.

These are but a few of the questions which present themselves when considering the mode and control of salt transport in the sweat duct. The answers are not easy to obtain: experiments *in vivo* are mostly not feasible and physiological studies *in vitro* on glands obtained by biopsy require exceedingly delicate techniques and sophisticated apparatus to deal with ducts one-tenth the diameter of a nephron. Much valuable groundwork has been done by K. Sato, using the monkey palm gland (*Rev. Physiol. Biochem. Pharmac.* 79, 51; 1977).

Now Quinton has answered the first question. He dissected out human sweat glands and cannulated 2-mm lengths of duct to perfuse them with solutions of varying composition, with or without ouabain. Measurement of the trans-epithelial potential difference enabled him to calculate the permeabilities of the tissue to Na⁺ and Cl⁻. The duct epithelium from cystic fibrosis sufferers turns out to have a permeability to Cl⁻ only about one-ninth of that of the normal tissue and this low value may be an underestimate. In fact, the ion is only slightly more permeable than SO₄²⁻, which is generally considered to be almost impermeable in epithelia. In these circumstances the transport of NaCl across the duct cell must be greatly impaired, which would account for the raised electrolyte level in the patients' sweat.

Quinton's work is a significant step forwards. It offers a challenge to membranologists to discover the reason for the low Cl⁻ permeability in the sweat duct and presumably in other affected glands in this inherited disease. It also raises the question of why a fundamental membrane defect, if such it is found to be, should be present in the exocrine glands and not in other tissues — the kidney, gut and red cells — where its presence would seem to be incompatible with normal function. □

V. Schwarz is Reader in Chemical Pathology at the Department of Biochemistry, University of Manchester, Oxford Road, Manchester M13 9PT.

Earthquake prediction

Electromagnetic emissions before earthquakes

from Chi-Yu King

LARGE earthquakes are frequently reported to be accompanied and preceded by some unusual phenomena, for example, deformation of the ground, changes in activity of small earthquakes, ground-water level and chemistry, and geomagnetic and electric fields, and strange behaviour of animals. These phenomena have been increasingly studied in recent years in many earthquake-prone countries, especially in China, Japan, US and USSR, with the intention of predicting future earthquakes. One such phenomenon is electromagnetic emission, some recent examples of which are reported by Gokhberg *et al.*¹ in a paper published in *Journal of Geophysical Research*.

Using a specially designed radio receiver located near Tokyo, Gokhberg *et al.* detected some anomalously high electromagnetic emissions on 31 March 1980 during a 30-min period before a magnitude 7 earthquake occurred 250 km away at a depth of 480 km. The anomalous emissions were recorded in two widely separated frequency bands (10–1,500 Hz and about 81 kHz). Similar anomalies were also recorded before two other earthquakes (magnitude 6 at a distance of 55 km and a depth of 75 km on 25 September 1980; magnitude 5 at a distance of 50 km and a depth of 60 km on 28 January 1981), as had been recorded previously in several different seismic areas of the USSR^{2,3}.

Reports of unusual electromagnetic emissions before earthquakes are also commonplace in several other countries, especially China⁴. A notable Chinese example is the magnitude 7.8 Tangshan earthquake on 28 July 1976 (focal depth 16 km). For 3 to 5 days before the earthquake, unusual interference was received by many military and civil radio-communication receivers in an area within about 250 km of Tangshan. In another case, before the Longling earthquakes (magnitudes 7.5 and 7.6) on 29 May 1976, rain-like noise was received by an ordinary household radio while 'earthquake light' was sighted⁴. The simultaneous observation of the light and the radio noise suggests that they both were caused by the same physical process, such as electrical discharge.

While some of the anomalous electromagnetic emissions, especially those reported by untrained observers, may conceivably be unreal or caused by some other geophysical or man-made processes (for example, distant lightning, meteors and arcing power lines), many such observations are well documented and the

existence of the phenomenon of earthquake-related electromagnetic emission must be seriously considered.

Several mechanisms have been proposed in the literature to explain the electromagnetic emissions, but none is considered satisfactory. One possible mechanism is charge separation and recombination produced by fractures of rock in the highly stressed focal volume of the impending earthquake. (It has been observed in the laboratory that rock specimens do generate impulsive electromagnetic emissions when being fractured^{5–7}.) However, it is not understood why the electromagnetic waves

thus produced may propagate many kilometres through the highly conductive Earth without being attenuated, especially for the deep earthquakes. Other proposed mechanisms invoke the piezoelectric effect of quartz-containing rocks^{8,9}, or streaming potentials produced by ground-water movement⁶. □

Chi-Yu King is in the US Geological Survey, Office of Earthquake Studies, Menlo Park, California 94025.

1. Gokhberg, M.B., Morgounov, V.A., Yoshino, T. & Timizawa, I. *J. geophys. Res.* **87**, 7824 (1982).
2. Sadovsky, M.A., Sobolev, G.A. & Migunov, N.I. *Dokl. Akad. Nauk USSR* **244** (2), 316 (1979).
3. Gokhberg, M.B., Morgounov, V.A. & Aronov, E.L. *Dokl. Akad. Nauk USSR* **248** (5), 1077 (1979).
4. Huang, L. & Deng, H. *Earthquake Light* (in Chinese) (Earthquake Press, Beijing, 1979).
5. Parkhomenko, E.I. & Vartyshev, Ju. N. *Physics of Earthquake Focus*, 51 (Nauka, Moscow, 1975).
6. Mizutani, H., Ishido, T., Yokokura, T. & Ohnishi, S. *Geophys. Res. Lett.* **3**, 365 (1976).
7. Nitsan, U. *Geophys. Res. Lett.* **4**, 333 (1977).
8. Finkelstein, D. & Powell, J.R. *Nature* **228**, 759 (1970).
9. Finkelstein, D., Hill, R.D. & Powell, J.R. *J. geophys. Res.* **78**, 992 (1973).

Cell membrane skeleton

The spread of spectrin

from Anthony J. Baines

THERE is probably a better understanding of the red cell membrane skeleton^{1,2} than of any other cytoskeletal structure and it is one of only two systems in which the interaction of actin with the membrane is understood at even a superficial level (the other being the intestinal brush border³). The red cell membrane skeleton confers stability on the lipid bilayer, which when deprived of this support, through loss of spectrin and actin, breaks up into small vesicles^{1,2}. These observations have suggested that even if the mechanical demands on the membranes of non-erythroid cells are less, they should contain structures somewhat analogous to membrane skeletons to maintain the stability of the plasma membranes and the membranes which bound subcellular organelles.

Actin has been demonstrated to be associated with the plasma membranes of many cell types. Recent attempts to identify the proteins involved in the actin-membrane interaction in non-erythroid cells have focused on a search for molecules structurally and functionally related to the structural proteins of the erythrocyte membrane — spectrin, 4.1, ankyrin and band 3. The use of anti-spectrin antibodies led to the identification of spectrin in all red cells so far examined⁴, but early results in the search for non-erythroid spectrins were not encouraging, for radioimmunoassay revealed no spectrin in several cultured cell lines⁵.

To qualify as a spectrin a protein must do more than simply react with anti-spectrin antibodies: it should have α and β subunits

with molecular weights in the region of 240,000 and 220,000 arranged in the cell as an ($\alpha\beta$)₂ tetramer; it should be able to cross-link and gel F-actin, with or without the participation of other proteins, and bind calmodulin and ankyrin; and it should be elongated and highly α -helical.

An axonally transported brain protein with subunit molecular weights 240,000 and 235,000, which had attracted interest because of its calmodulin- and actin-binding properties, was purified by several groups working independently^{6–8, 29}; other proteins of apparently identical properties were identified by their cross-reaction with spectrin in brain⁹ and chicken cardiac myocytes¹⁰. Various names^{7,9,29} have been proposed for the brain protein, but it is fodrin⁶ (from the Greek *fodros*, a lining) that has stuck, derived from the observation that, as judged by immunofluorescence microscopy, it lines the cytoplasmic face of nerve cell plasma membranes⁶.

Weber and co-workers¹¹ have purified a different protein from intestinal brush border which cross-reacts with spectrin. Its subunit molecular weights are 240,000 and 260,000 respectively, for the counterparts of the α and β chains of spectrin, and it has been dubbed TW 240/260.

Several reports have characterized fodrin and TW 240/260 in great detail and all seem in remarkable agreement that they are genuinely spectrinesque; they are both ($\alpha\beta$)₂ tetramers^{9,11,29} with morphologies similar to that of spectrin^{9,11,29} and can cross-link and gel F-actin^{9,12,13}. The salt dependence of fodrin's sedimentation

characteristics is strongly reminiscent of spectrin¹³ and it binds to red cell ankyrin^{9,13}, the protein that attaches spectrin to the membrane.

The α subunit of these proteins is evidently the common denominator¹⁴, certainly in terms of its size, strong antigenic cross-reaction and ability to bind calmodulin¹⁴. The extent of the sequence homology is uncertain, for peptide fingerprints generated by partial proteolysis¹⁴ show considerable homology whereas limit digests reveal very little⁹; this, however, is also true of pig and human spectrin⁹.

The α subunit may be a widespread unit of structure as similar antigens can be detected in such diverse tissues as lens, liver, cardiac and skeletal muscle in addition to nervous tissues and some cultured cell lines¹⁵.

Antigenic cross-reactivity can be a treacherous criterion if too incautiously interpreted. Davis and Bennett¹⁶ found that microtubule-associated protein 2 (MAP 2) stained with anti-spectrin antibody on immunoblots of SDS gels, and indeed MAP 2 is a high-molecular-weight actin-binding protein¹⁷ with a flexible rod morphology¹⁸. However, the extent of antigenic cross-reactivity is limited¹⁶ (the relationship seems again to be with α -spectrin¹⁶), and MAP 2 is reported to have an α -helix content of only 4 per cent¹⁸ compared with some 75 per cent for spectrin¹. Moreover, metal-shadowed preparations show a protein that is monomeric with respect to subunit composition¹⁸. This protein probably belongs to a different structural category, but it would be interesting to establish where the spectrin-related element resides, since it might provide a marker for an actin-binding domain in spectrin.

Filamin too, because of its size¹⁹, morphology²⁰ and ability to cross-link F-actin¹⁹, was conjectured to be a non-erythroid spectrin, but it does not bind ankyrin²⁰ or calmodulin¹⁴, nor will spectrin bind anti-filamin antibody¹⁹ and peptide fingerprints obtained by partial proteolysis bear no resemblance to those of spectrin²¹. The comparison would therefore seem as specious as that made at an earlier stage between spectrin and myosin.

In none of the stages of the purification of fodrin^{6,9} or of TW 240/260 (ref. 11) is there any sign of a protein that could be the counterpart of 4.1, the molecule which modulates the interaction of spectrin and actin in the red cell. However, this is not to say that such a protein does not exist: two groups^{22,23} have found proteins cross-reacting with 4.1 in cells such as granulocytes, platelets and fibroblasts. The fibroblast protein is located close to stress fibres and has a molecular weight similar to that of 4.1 (ref. 23).

As fodrin is located close to the plasma membrane, the existence of a membrane-anchoring system comprising ankyrin and band 3 or their counterparts is predicted.

Ankyrin has been found in many cells and tissues by radioimmunoassay²⁴, and proteins within the expected range of molecular weight and reacting with anti-ankyrin and anti-band 3 antibodies have recently been found in the plasma membrane of liver cells together with a fodrin²⁵. Burridge²⁶ has also reported the purification of a protein of the same molecular weight as ankyrin from chicken gizzard; immunofluorescence localizes a cross-reacting protein in the focal adhesion plaques of HeLa cells, which is where one might hope to find an ankyrin.

Ankyrin also bears a considerable resemblance to MAP 1 (ref. 27), with which it cross-reacts antigenically. Moreover, it co-purifies with microtubules during successive cycles of polymerization and depolymerization; which is interesting, considering especially that another microtubule-associated protein (MAP 2) is immunologically related to spectrin. Although fodrin does not interact with microtubules, its possession of an ankyrin binding site and the apparent relatedness of MAP 1 to ankyrin invites the conjecture that perhaps fodrin could be associated with microtubules via MAP 1: as fodrin is derived from brain plasma membranes there must be a binding site (presumably ankyrin-like) on the membranes for fodrin. This opens up possibilities for some wonderfully roccoco arrangements of microtubules, microfilaments and membranes.

In his review of the red cell membrane skeleton only three years ago, Lux²⁸ commented that it was not at all clear how non-erythroid cells could survive without membrane skeletons, there being at that time no indication that they had such structures. The evidence that they have bears on more than just membrane structure. Spectrin has been implicated in the restriction of lateral movement of red cell surface receptors and although the

diffusion coefficients of cell-surface receptors are generally higher in non-erythroid cells, they are still lower than that which corresponds to free diffusion and, in particular, the receptors on fibroblast surfaces move faster in the direction parallel to the stress fibres than in the perpendicular. □

Anthony J. Baines is at the Department of Biophysics, King's College, 26-29 Drury Lane, London WC2B 5RL.

1. Gratzner, W.B. *Biochem. J.* **198**, 1 (1981).
2. Branton, D., Cohen, C.M. & Tyler, J. *Cell* **24**, (1981).
3. Weber, K. & Osborne, M. in *Cellular Controls in Differentiation* (eds Lloyd, C.W. & Rees, D.A.) (Academic, London, 1981).
4. Pinder, J.C., Phetean, J. & Gratzner, W.B. *FEBS Lett.* **92**, 278 (1978).
5. Hiller, G. & Weber, K. *Nature* **266**, 181 (1977).
6. Levine, J. & Willard, M. *J. Cell Biol.* **90**, 631 (1981).
7. Shimo-Oka, T. & Watanabe, Y. *J. Biochem.* **90**, 1287 (1981).
8. Davis, P.J.A. & Klee, C.B. *Biochem. Int.* **3**, 203 (1981).
9. Bennett, V., Davis, J. & Fowler, W.E. *Nature* **299**, 126 (1982).
10. Goodman, S.R., Zagon, I.S. & Kulikowski, R.R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7570 (1981).
11. Glenny, J.R., Glenny, P., Osborne, M. & Weber, K. *Cell* **28**, 843 (1982).
12. Glenny, J.R., Glenny, P. & Weber, K. *J. biol. Chem.* **257**, 9781 (1982).
13. Burridge, K., Kelly, T., Mangeat, P. *J. Cell Biol.* **95**, 478 (1982).
14. Glenny, J.R., Glenny, P. & Weber, K. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4002 (1982).
15. Repasky, E.A., Granger, B.L. & Lazarides, E. *Cell* **29**, 821 (1982).
16. Davis, J. & Bennett, V. *J. biol. Chem.* **257**, 5816 (1982).
17. Sattilano, R.F., Dentler, W.L. & LeCluyse, E.L. *J. Cell Biol.* **90**, 467 (1981).
18. Voter, W.A. & Erikson, H.P. *J. ultrastruc. Res.* **80**, 374 (1982).
19. Wallach, D., Davis, P.J.A. & Pastan, I. *J. biol. Chem.* **253**, 3328 (1978).
20. Tyler, J.M., Anderson, J.M. & Branton, D. *J. Cell Biol.* **85**, 489 (1980).
21. Calvert, R. & Gratzner, W.B. *FEBS Lett.* **86**, 247 (1978).
22. Spiegel, J.E., Beardsley, D.S. & Lux, S.E. *Fedn. Proc.* **41**, 657 (1982).
23. Cohen, C.M., Foley, S.F. & Korsgren, C. *Nature* **299**, 648 (1982).
24. Bennett, V. *Nature* **281**, 597 (1978).
25. Bennett, V., Davis, J. & Fowler, W.E. *Proc. R. Soc. B.* (in the press).
26. Burridge, K., Kelly, T. & Campbell, L. *Proc. R. Soc. B.* (in the press).
27. Bennett, V. & Davis, J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7550 (1981).
28. Lux, S.E. *Nature* **281**, 426 (1979).
29. Kakiuchi, S. *et al. Biomed. Res.* **3**, 400 (1982).



100 years ago

SCIENTIFIC HERESIES IN CHINA

THE belief of the old writer who explained the disappearance of the swallows in autumn, by the assumption that they all rolled themselves together into a mass, beak to beak and wing to wing, and plunged "fluminibus lacubusque," there to remain until the return of spring, is but an instance of the thousand and one theories which have from all time been held by unscientific observers on the subject of the annual appearance and disappearance of migrating birds. It is not so very long ago that even among ourselves the question of the migration of sand martens was a moot point;

and it need not be a matter of surprise, therefore, that the Chinese, though keen observers of nature, should be guilty of holding heretical hypotheses to account for the presence in summer and the absence in winter of birds of passage, as well as for other processes of nature which are enacted before them. Like most common fallacies, those current in China on these subjects are derived from ancient authorities, but, unfortunately, in the case of the Chinese, whose respect for everything that is old is supreme, this antiquity only entails upon them the more unquestioning faith. They are, therefore, perfectly content to believe that the disappearance of quails in autumn is sufficiently accounted for by the assumption that at that time of the year they are transformed into moles, and that in spring they succeed in reappearing again in beaks and feathers. The explanation of this fallacy is simple enough. The ploughman who in the spring and summer has seen the quails flitting among the mole-hills, finds that when, the birds having disappeared, he ploughs over the same land in winter, the moles, which he has not before seen, are the sole occupants of the ground.

From *Nature* **27**, 342, 8 February 1883.

The ecology of carnivore social behaviour

David W. Macdonald

Animal Behaviour Research Group, Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

Diverse selective pressures have contributed to the evolution of the varied social groups of carnivores: the benefits of strength of numbers for defence of kills and territory, and in the hunting and killing of large prey; the ability to intimidate predators and to be vigilant against their approaches; the potential for information transfer and social learning, and a suite of alloparental behaviour patterns. Each of these may operate within the constraints upon group size and home range size set by patterns of resource dispersion. Between and within species, the magnitudes of costs and benefits attendant upon group life vary with circumstances and between individuals.

TRADITIONALLY, two selective pressures have been invoked to explain why some of the Carnivora live in groups: the need for assistance in hunting and killing large prey, and the need for defence against attacks by other predators¹⁻⁴. Wolves, for example, hunt cooperatively to capture their large, and sometimes dangerous, quarry while, in contrast, red foxes, *Vulpes vulpes*, kill small rodents by stealth and do better to hunt alone^{5,6}. Representatives of at least four genera of mongooses benefit from living in packs through the increased vigilance provided against marauding predators, and through the potential for intimidating predators^{1,2,7} by their group defiance.

Neither of the above selective pressures explains, however, why some species live in groups but travel and hunt alone. Clans of European badger, *Meles meles*, living in communal setts^{8,9}, parties of giant otter, *Pteronura brasiliensis*¹⁰, groups of red foxes¹¹, colonies of 'feral' cats (refs 12, 13 and D.W.M., P. J. Apps and G. M. Carr, in preparation), groups of palm civets, *Nandinia binotata*¹⁴, and groups of brown hyaena, *Hyaena brunnea*¹⁵, serve as examples from five families where other reasons must be sought to explain social grouping. Study of these carnivores indicates that selective pressures previously thought of as rather incidental, have given rise to complex social groupings. Benefits, which vary according to the species or population, come not only from cooperative hunting, defence of territory and prey, and defence against predators but also from the opportunities to learn from other group members' experiences, division of labour, care of the sick and several advantages associated with alloparental behaviour, that is, the care of the offspring of others. However, all the many benefits of group living must operate within a framework of constraints determined largely by the dispersion and abundance of available resources, especially food. Considering the great diversity of carnivore social systems, both inter- and intraspecific variation in the ecological costs of sociality is as relevant as is variation in the behavioural benefits. Here I shall argue that there are ecological circumstances in which the benefits of 'spatial groups' (and equally perhaps the evolutionary origins of contemporary highly social groups) may have little or nothing to do with advantages directly consequent on sociality. Rather, resource (particularly food) dispersion is fundamental to the spacing and structure of carnivore society in that it may set the limits to the group and territory sizes within which other combinations of selective pressures operate.

Resource dispersion hypothesis

The idea that certain patterns in the availability of food (and other resources) may allow group living, originates largely from work on birds¹⁶⁻²⁰. The general principle applies to all social systems, and has been applied in particular to several mammalian societies (bats²¹, primates²²⁻²⁴, antelope²⁵), several carnivores^{8,9,11,26-30}, and generally to mammalian mating systems^{20,31}, and, very convincingly, to the social organization of badgers³⁰.

Two extreme examples show how the pattern of food availability influences the way in which carnivore groupings develop. In Portsmouth, female dockyard cats live in groups (mean 5.4 individuals) sharing large areas of their individual home ranges (mean 1.8 hectares)¹³. Each colony's range is centred on a 'skip' of refuse containing edible offal and from which other female cats are debarred. Similarly, golden jackals, *Canis aureus*, observed in Israel form stable social groups of as many as 25 adults around a feeding site³². These jackals ring their territory with communal latrines, patrol the borders as a cohesive group, share food and repel intruders collaboratively—behaviour very different from that found in other ecological circumstances. These cats and jackals fed largely from one or a few sites whose existence provided the possibility of group formation and whose collaborative defence provided one obvious advantage of group living.

The factors permitting and promoting group formation are less clear for some other species, such as red foxes. In rural suburbia outside Oxford¹¹, groups of red foxes consist of one male and up to five adult vixens (mean 3.4 females, $n = 5$). Each individual is in physical contact with another for as little as 4 min each night³³, yet they are social units in that members have clearcut, amicable relationships and share a fairly well defined group territory (mean 45.2 hectares, $n = 7$). Their most important foods are earthworms, scraps scavenged from human settlements and fruits^{11,28}. These foods vary in their spatial and temporal availability. Worms are caught on the surface³⁴ and only emerge when the microclimate is appropriate. Local microclimates vary independently, and thus from night to night, and even within one night, patches of higher availability of worms occur in different parts of each fox's territory. Similarly, a garden may be rich in household scraps one night but devoid of them the next, and fruits are seasonal. The minimum territory size required to sustain a pair of foxes, and the territory's configuration, may thus be determined by the dispersion of transient patches of available prey (ephemeral from night to night, or shifting from one season to the next), and constrained by the costs of defence of irregularly shaped areas. In contrast, group sizes are limited by the availability of prey in these patches. In this context, the resource dispersion hypothesis (see refs 11, 15, 29, 30) suggests that the smallest home range with an economically defensible³⁵ configuration which will reliably support a pair of red foxes (on a bad night or a bad year) may sometimes support additional foxes. These other foxes are tolerated in numbers and at times when any costs (to the basic pair) due to their presence are outweighed by the overall benefits. If the costs are small then of course the advantages need not be great; thus, for example, a minimal advantage of group membership might be no more than the avoidance of dispersal³⁶ (the surrounding habitat may be saturated or dangerous) or guaranteed access to a den site, water hole or familiar hunting ground. Such benefits have little (or even nothing) to do with social behaviour *per se* but need not result

in groups that are merely structureless aggregations. Just as territories can usefully be defined as ranges which overlap less than expected by chance³⁷, so members of 'spatial groups' can be defined as individuals whose ranges overlap more than expected by chance. Members of spatial groups are likely to meet frequently and would be readily available as beneficiaries of advantages derived more directly from sociality. In the case of red foxes these social advantages include alloparental behaviour^{36,38}, increased boundary surveillance and occasional corporate attack of trespassers²⁸.

Resources, group size and territory size

From the resource dispersion hypothesis one would not necessarily expect (even in broadly similar habitats) any relationship between group size and territory area, as the two are argued to be affected, largely independently, by the abundance and dispersion of available food respectively. However, for a given patch-richness one would expect minimum sized territories to be larger where patches were more dispersed (Fig. 1). The resource dispersion hypothesis neither discounts nor precludes the possibility that animals may strive to maintain territories and/or groups that are larger than the minimum^{38,39}, rather it provides an ecological explanation of how, at minimum costs to themselves, members of a pair could incorporate additional group members into even the smallest economically defensible territory which will sustain them. The absolutely greater (but proportionately less) costs of border defence for the territorial expansion required to support larger groups might be offset by greater benefits. On the other hand, advantages to a carnivore of knowing its home range extremely well may favour ranges which are as large as necessary but as small as possible.

To measure the availability of resources, such as prey, is notoriously difficult. However, to the extent that habitat types are an index of prey abundance (for example, a ploughed field supports fewer mice than does mature grassland), the geometry of the habitats within a territory may be reflected in the pattern of prey availability. In this case, the sizes and configurations of territories will be such as to encompass at least a minimum total area of key habitats, representing rich prey-patches scattered in otherwise barren areas. Studies of three species of carnivore illustrate the argument. The European badger, *Meles meles*, specializes on one prey, the earthworm^{8,9}, whose abundance is readily measured and whose availability in given habitats can be estimated. This enabled Kruuk and Parish³⁰ to show that variations in badger society reflect resource dispersion. Badgers forage in patches of short grass pasture which have high earthworm availability, 4–5 areas of which are found per territory. If the patches are widely dispersed, territories are correspondingly larger, but there is no correlation between territory size and group size. In contrast, mean 'patch quality' (that is, abundance of earthworms) in each territory did correlate with group size. Thus, patch dispersion influences territory size, whereas patch quality independently influences group size. The problem remains of what social mechanism adjusts group sizes to patch quality.

In one study of red foxes¹¹, adjacent territories varied in size between 18.6 and 72.0 hectares and groups consisted of 3–6 adults, but larger groups did not occupy larger territories. Radio tracking, diet analysis and prey sampling indicated that residential habitats (houses, gardens and orchards) were intensively used and rich in food. In adjacent arable farmland prey was scarce. Analysis of seven adjacent territories showed a comparable area of residential habitat in each [mean = 10.4 ± 3.5 hectares ($\pm$ s.d.)] and territory configurations clearly designed to encompass these key patches. The most constant component in each territory was the number of houses (mean = 23.7 ± 4.8) (probably an indicator of the availability of scraps).

Three groups of Arctic foxes, *Alopex lagopus*, studied in the fjords of north-west Iceland lived in units of one male and two adult vixens²⁹. The groups occupied territories varying between 8.6 and 18.5 km², and gained 60–80% of their food by beach-

combing. Carrion could be found only in those 'productive' bays favoured by the drift of the current. Despite their different areas and total lengths of coastline, each territory encompassed rather constant lengths of 'productive' coastline. More strikingly, each territory's coast yielded annually a comparable bulk of driftwood to local farmers (mean = $1,900 \pm 44$ fence posts), and so presumably also a comparable availability of carrion to foxes.

Territory sizes in contrasting habitats

Even within broadly similar habitats there is considerable intraspecific variation in the basic parameters of carnivore social organization—that is, group size and home range size (equivalent to territory for most species mentioned here). Between populations in contrasting habitats the variation can be enormous: in the uniform and sparse habitats of the North American prairies, red foxes live in pairs within large (500–2,000 hectares) territories⁴⁰; in the diverse, rich habitats on the outskirts of Oxford they form groups within small (10–100 hectares) territories¹¹, while in the taiga forests of Sweden where prey abundance cycles from year to year, territory sizes of 700 hectares remain unchanged in area (perhaps adjusted to the worst years) while group sizes vary with vole abundance³⁹. Like those between adjacent territories, these variations between dramatically different habitats seem explicable in terms of variation in the constraints imposed by food dispersion on group formation: where prey are more homogeneously distributed, as in intensive arable land of the prairies, then the smallest territory that will support a pair may not support another adult. Where there are cycles of prey availability between years, as in the taiga, then the smallest territory which supports a pair during troughs of vole abundance can incorporate additional adults during the peaks of vole numbers. During bottleneck years or seasons, groups shrink, but territories do not necessarily contract³⁹. The regulation of group size involves social status: there is evidence that amongst wolves^{41,98} and foxes²⁸ low ranking animals are expelled as aggression mounts with diminishing food availability. These observations are compatible with Bertram's²⁷ view of lion pride and range sizes—in the long term, pride ranges probably respond to changes in food availability, but in the shorter term they remain constant, adapted perhaps to the worst of recent conditions, but with the option of altering pride size through recruitment of sub-adult females. That social factors affect the regulation of group size among many carnivores^{42–44} is of fundamental importance to their population dynamics^{16,45}.

Females as a resource

Female, rather than food, dispersion may underline spacing of males in those species where several exclusive female territories are encompassed within that of one male³¹ (for example, bobcat, *Felis rufus*⁴⁶, stoat, *Mustela erminea*⁴⁷, and wolverine, *Gulo gulo* (refs 48, 49 and A. Magoun, personal communication)). Amongst both Mustelidae and Ursidae there is a spectrum from male monopoly of several female territories and greatest sexual dimorphism (100%) and carnivory, to species where the sexes are equally sized, frugivores or omnivores inclined to more amicable social ties involving monogamy or group formation, such as in Tayra, *Tayra barbarosa* or sloth bears, *Melursus ursinus*, and sun bears, *Helarctos malayensis* (10–25% dimorphism) (ref. 50 and F. Bunnell, personal communication). As work on supposedly less sociable carnivore species continues⁵¹, the pattern in which the spatial organization of females is determined by the distribution of food resources, and that of males is determined by the distribution of females, may turn out to be widespread.

Group size and hunting success

Cooperative hunting has only been documented for members of the Hyaenidae, Canidae and Felidae of the seven families of carnivores (although some members of all families forage in bands). Among the canids, the hunting dogs, *Lycaon pictus*,

almost invariably hunt cooperatively⁵², as do dholes, *Cuon alpinus*⁵³, but most other pack hunters are only facultatively cooperative in their pursuit of prey. For example, wolves⁵⁴, coyotes, *Canis latrans*⁵⁵, dingos, *Canis dingo* (L. C. Corbett, personal communication), jackals, *C. aureus* and *Canis mesomelas*^{44,56}, and probably also bush dogs, *Speothos venaticus*, regularly, but not invariably, hunt together for large prey, whereas other species do so infrequently. Bat-eared foxes, *Otocyon megalotis*, are unusual among canids in foraging in small family parties but for invertebrate prey^{57,58}.

Members of a clan of spotted hyaena, *Crocuta crocuta*, may assemble into temporary hunting parties, of which larger parties (mean = 10.8 members) hunt zebra, *Equus burchelli*, smaller ones (mean = 2.5) hunt adult wildebeest, *Connochaetes taurinus*, and yet smaller ones (mean = 1.2) hunt the fawns of gazelle, *Gazella thomsoni*⁶. Parties of a given size search for 'their' prey species, ignoring others and thus indicate a remarkable corporate decision about which prey to hunt, even before sensing their victims. However, although there is a positive correlation between prey size and social group size across the Carnivora⁵⁹, the variance in hunting party membership is only partly explained by the increased difficulty of killing larger, fleet prey. Also, jackals cooperate to hunt small prey⁵⁶, whereas kit foxes, *Vulpes macrotis*, hunt alone for rabbits twice their weight⁶⁰. The question thus arises of whether hunting together functions solely to increase efficiency of hunting and killing large prey.

Clearly, cooperative hunting is not the sole benefit of group living for, in at least some cooperative hunters, groups may be much larger than hunting parties (for example, spotted hyaenas⁶ and lions⁶¹). Even the size of hunting parties is probably not primarily governed by the requirements of cooperative killing^{62,63}. Nevertheless, in some circumstances cooperative hunting is more efficient than foraging alone: two silver-backed jackals, *Canis mesomelas*, can thwart a mother gazelle's defence of her fawn whereas a single jackal can rarely do so⁵⁶. Similarly spotted hyaenas hunting in pairs can invariably separate a wildebeest calf from its mother⁶. Larger packs of coyotes may kill relatively more larger prey such as deer⁵⁵, and van Orsdol⁶¹ reports that hunting efficiency increases with the size of lion hunting parties up to a maximum of four or five members. The defence of kills from conspecifics and scavengers⁶⁴ may also encourage group hunting and this selective pressure may affect different species differently: for example Lamprecht⁶⁴ has argued that jackals do not hunt in larger groups for bigger prey because even their corporate power would be insufficient to repel large scavengers. He also argued⁶² that other (larger) species hunt in larger groups primarily to secure the collective strength to defend their spoils.

I suggest that all the possible advantages of group hunting act within the fundamental framework of constraints set by the dispersal of resources, such that an imaginary, smaller group will require a territory just as large as that occupied by the observed larger group. As a speculative example, lions often ambush their prey around water holes⁶⁵. An ancestral lion might have required several water holes to ensure access to prey across seasons and through ungulate migrations. However, each waterhole, when productive at all, might easily support several lions and some or all of the combined advantages of increased hunting success, nursing coalitions, defence of kills, the need to avoid interference with each other's hunting, could then favour cooperative hunting as well as group living. I therefore argue that the evolutionary catalyst for group living (that is, prey dispersion) could have been the same for, say, red fox and wolf and that the contemporary nature of the groups (for example, solitary or collaborative hunting) differs partly due to the prey's size and ability to defend itself.

That several hunting units of variable composition operate within lion prides (and spotted hyaena clans⁶) and that any advantage of cooperative hunting only applies to hunting parties of 4–5 lions⁵⁹ whereas prides may number up to 37, and even hunting parties may number over 4–5, indicates that factors

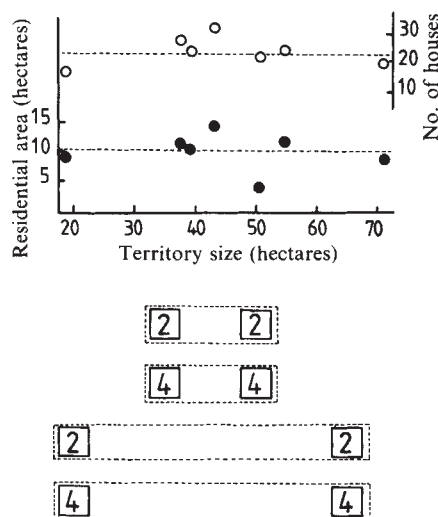


Fig. 1 A simplified case of the resource dispersion hypothesis: temporal variation in the availability of prey means that even a pair of animals require two patches (symbolized as squares) and the dispersion of these patches determines the minimum necessary economically defensible area, whereas patch richness (for example, values of 2 versus 4 indicate patches which support either two or four animals) determines the maximum group size that can be supported in that minimum area. In a habitat where dispersal was disadvantageous it could thus be advantageous for these four animals to form a spatial group purely because of the availability of resources and irrespective of any direct benefits of sociability *per se*. These schematic patches may be thought of as gardens with nightly variations in the availability of scraps for foxes¹¹, or pastures that are more or less bountiful for foxes and badgers foraging for earthworms depending on prevailing wind direction^{8,9,30,34}, or they could be asynchronously fruiting trees upon which frugivores feed^{22,23}. The graph shows that red fox territories of widely differing sizes embrace similar areas of 'residential' habitat (which is prime foraging habitat) and roughly the same number of human residences from which food was scavenged from bird tables and compost heaps, and on whose lawns earthworms were caught. Territorial configurations were clearly adapted to embrace clumps of houses and gardens¹¹.

other than the collaborative hunting of even the largest prey underlie group living and hunting by cooperative hunters today. Schaller's⁶⁵ lions only doubled their success by hunting in pairs but two or more lions were better able to repel scavenging hyaenas after making a kill. Packs of hunting dogs contain 1–18 members; some packs learn, from pack tradition, to kill zebra using a restraining nose-hold and a pack of four and one of eight seemed equally adept at this^{66–69}. Similarly, two or three dholes can overpower a 60 kg deer apparently as well as a pack of the average (8.3 adults) or maximum (16–20) size⁵³.

Population differences in cooperative hunting

In Texas, coyotes feed mainly on rodents⁷⁰, but they nevertheless live in groups, whereas on a similar diet, coyotes in the Rocky Mountains⁷¹ live in pairs. Elsewhere, coyotes live in packs and feed almost exclusively on ungulate carrion⁷² which larger packs are more successful at defending from rival packs than are pairs^{55,71}. Also, larger packs appear to be more effective hunters of deer than are smaller ones⁵⁵. Thus, a diet of large prey may favour (but not necessitate) the formation of large hunting parties, and the fact that groups and hunting parties are often apparently supra-optimal for hunting presumably means that their membership is in response to other selective pressures. In coyotes, hunting by man may depress group sizes both directly and by creating vacant territories and so reducing the costs of dispersal.

Territory size versus group size

Considering adjacent territories in broadly similar habitats, no correlation between group size and territory size has been found

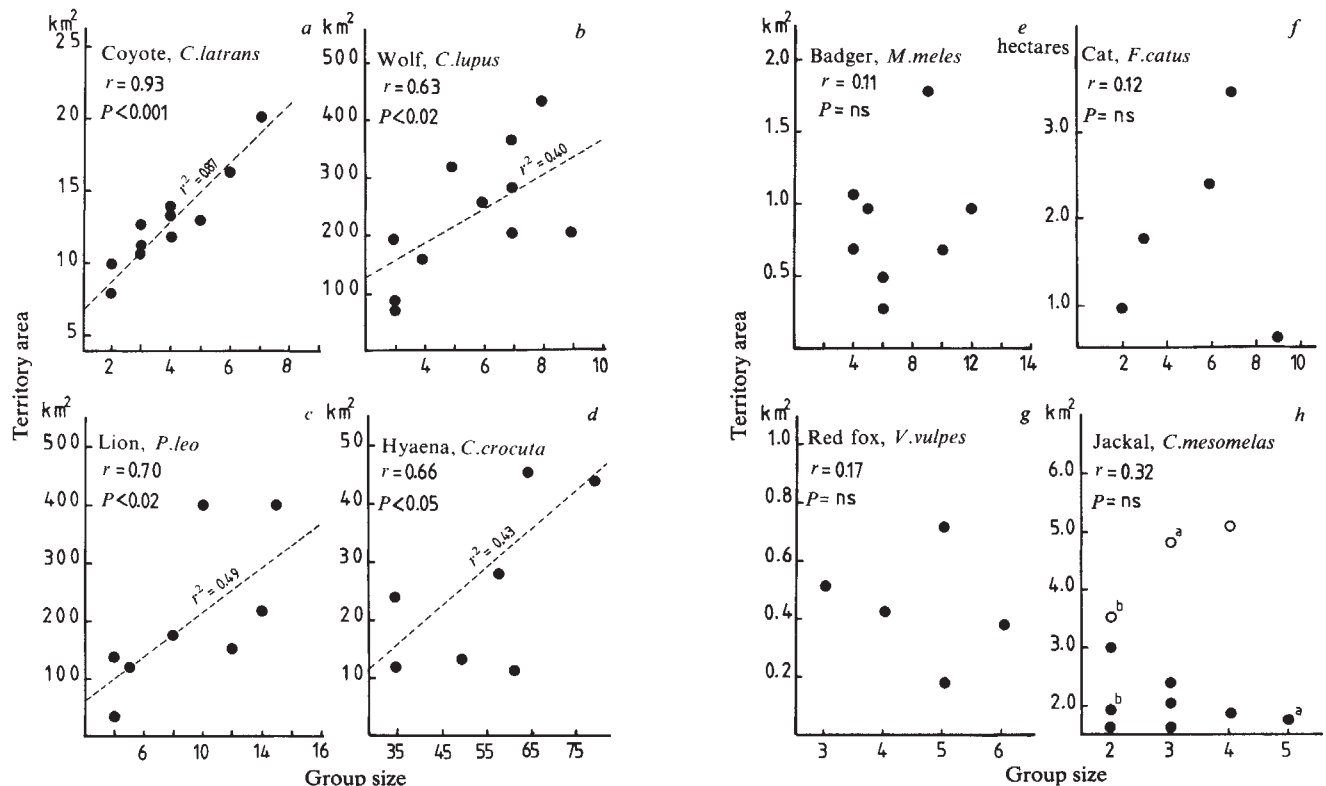


Fig. 2 Two patterns emerge from plotting group sizes against territory areas for eight species of carnivore, representing four families. In *a-d* there are variously precise correlations between the two parameters, whereas there is no such relationship for the species in *e-h*. NS, not significant. Each graph concerns more or less adjacent ranges from broadly similar habitats (which minimizes the effect of confounding variables, such as any inter-population variation in body size which might affect range sizes). Territory sizes were measured from maps, using a planimeter, and combined with data from tables in refs 6, 8, 9, 11, 13, 44, 55, 65, 94 and personal communications from the same authors. (In each of these, and many other carnivore species, home range is the same as territory.) The resource dispersion hypothesis can explain, but does not necessitate, the absence of correlation between territory and group sizes. In *b*, group sizes refer to adult wolves in spring. In *h*, open circles represent jackal territory and group sizes in 1 yr of prey scarcity; these were excluded from the analysis (a and b were jackal territories studied in 2 yr)⁴⁴. These varying relationships within populations do not detract from (and are actually incorporated in) the broader interspecific correlation which Gittleman and Harvey⁹⁷ found between home range sizes and metabolic needs (a measure which includes group sizes) across the carnivores. Presumably, the selective pressures which give rise to interspecific variation are the same as those causing intraspecific variation.

among several group-living but solitarily hunting species (Fig. 2b). In contrast, membership of packs of several cooperatively hunting species does increase with territory size (Fig. 2a). Pack size accounts for 87% of the variance in home range size amongst one coyote population, and in two cases range sizes were reduced after the death of pack members⁵⁵. Bygott *et al.*⁷³ report how a large coalition of male lions may successively gain control of two or even three formerly independent territories and hence prides of lionesses. Male cheetah, *Acinonyx jubatus*, form similar coalitions of 2-4 members⁶⁹, and tigers, *Panthera tigris*, expand their territories into that of a sick neighbour⁷⁴. In some areas, red fox territories may expand after the death of a neighbour⁷⁵, and elsewhere³⁹ territory sizes remain the same in years of widely different food abundance. Thus carnivores sometimes defend a larger area than the minimum necessary for a pair, and sometimes they thus incorporate additional group members. The advantages of expansionism to members of male felid coalitions include procuring more matings and surviving serious fights^{69,73}, whereas the coyotes probably found and retained more moose carcasses⁵⁵.

The differences between cooperative and solitary hunters in Fig. 2 are puzzling: one would expect that by travelling together, larger parties of cooperative hunters were better placed to overwhelm their neighbours and hence to secure larger territories. However, packs of banded mongoose travel together, forage individually and show no correlation between group and pack sizes, although larger groups do win fights (ref. 7 and J. Rood, personal communication). The situation is complicated as although all four species in Fig. 2b are cooperative hunters, they differ with regard to the relationship between their social

and foraging units. Each may enjoy a wide range of benefits from group living, but there is little evidence of their comparative importance. Some benefits to group living are mentioned below.

Anti-predator behaviour

Rood⁷ found that baby-sitting by adults reduced predation on young dwarf mongooses by slender and banded mongooses; he also described (personal communication) a dramatic rescue from an eagle's talons of a captured banded mongoose by the alpha male of the pack. Predators are more likely to be spotted and/or intimidated by larger groups of prey⁷⁶, and this advantage has been ascribed especially to social viverrids^{1,2}. Vigilance for predators may be an important benefit of cohesive group travel, and, as Rasa¹ argues, where life in open terrain puts small carnivores at particular risk (indeed, across the order group living is associated with open habitat⁵⁹).

Kinship

It is axiomatic that groups may be expected to develop where the costs of the resident's fitness of evicting an interloper exceed those of tolerating it. At worst, an extra 'group member' could thus be a harmless but stubborn sitting tenant. However, to an extent which depends on their degree of relationship⁷⁷, the net cost of tolerance to the occupier's fitness is diminished in proportion to the benefit accruing to the interloper, if they are related. In most carnivore species for which there is evidence, groups consist at least largely of relatives^{78,82} (for example, lions⁷⁹, wolves^{41,43,98}, hunting dogs^{66,67}, spotted and brown

hyaena^{6,15}, two species of jackal^{32,44}, two species of fox^{29,36}, two of mongoose⁷, coyotes (refs 55, 72), dingo³⁶ and dholes⁵⁴; coatis, *Nasua narica*, provide the only documented exception where female group membership (and alloparental behaviour) is not largely restricted to kin⁸⁰. Male lions increase their chances of procuring and retaining a pride if they form coalitions^{73,79} and 58% of coalitions comprise related males⁸¹. All else being equal, kin selection should favour joining forces with a relative⁷⁹, but the benefit of kinship is merely a bonus to the overwhelming advantage of any coalition. An effective unrelated companion is presumably preferable to an ineffective related one. Irrespective of relatedness, male lions squabble within a coalition⁸¹, the winners securing more key matings, emphasizing that while it may be better to lose to a relative, it is best not to lose at all.

Alloparental behaviour

In carnivore societies, the guarding and defence of and provision for young other than one's own is common⁷⁸. Among feral cats, immature males baby-sit kittens^{12,13}, and breeding females guard, nurse and feed each other's offspring, as do lionesses⁷⁹. Similar collaborative rearing is found among many canids, among which it is also common for non-breeding animals to participate as alloparents^{29,36,44,53,54,72,83}. (Among dwarf mongooses, collaboration extends to care of sick adults⁸⁴.) Information is less complete for other families, but one interesting case is the brown hyaena, among whose groups only dominant females breed, and non-breeders of both sexes provide for the young. The young are sired by nomadic males which have only fleeting associations with the groups whose females they fertilize⁸⁵ and whose cuckolded males act as helpers. Of the Viverridae, *Mungos mungo* and *Helogale parvula* both show communal care of young, but several females breed in *Mungos* packs, whereas only one does in those of *Helogale*⁷. Subordinate members of some carnivore groups (especially canids) endure at least temporary reproductive suppression. When a subordinate does give birth her offspring may be neglected³⁶, killed⁸⁶, or stolen⁶⁹, although on other occasions females may den communally^{36,72,83} and cooperate.

Are 'helpers' helpful?

Although the term 'helper' is often used as a synonym for a non-breeding group member, proof of their helpfulness is scant. Alloparents increase the survivorship of silver-backed and, probably, golden jackals⁴⁴; they probably sometimes contribute to the survival of red fox cubs³⁶, and may adopt orphans³⁸. One would expect the role of potential helpers to be tempered by prevailing ecological circumstances⁷⁸ and this may explain why yearling hunting dogs^{66,67} help feed pups when food is plentiful, but steal from them when food is scarce. Similarly, pup survival for Minnisotian wolves was slightly reduced among bigger packs (with more potential alloparents) during years when the wolves were known to be under mounting population pressure⁸⁷. As most social groups consist of relatives, most alloparents may benefit from kinship⁷⁷, mutualism and reciprocity^{88,89} amongst other possibilities. However, unrelated immigrants occur in dwarf mongoose packs, and one newcomer was a more assiduous helper than were some original pack members⁷. These immigrant helpers benefited by breeding earlier than their counterparts who stayed at home.

Flexible social behaviour

The great flexibility of carnivore societies is perhaps most conspicuous regarding territory and group sizes—wolf territories vary between 50 and 1,000 km² and contain packs of 2–20 members^{34,90}, and red fox territories vary 70-fold in area¹¹ and can vary 600-fold from those of otherwise similar Arctic foxes²⁹. Similarly, raccoon territories vary from 0.5 to 50 km² (D. R. Voigt, personal communication) and those of grizzly bear from 24 to 1,054 km² (F. Bunnell, personal communication). The average size of groups of European badgers varies between habitats, from 3.0 to 7.6 members and average terri-

tory sizes span 22 to over 200 hectares^{30,91}. Interspecifically, the wolverine is less than twice the weight of the American badger⁹², *Taxidea taxus*, yet its home range can be 400-fold larger (both are mustelids). Doubtless there is similarly extensive, but less conspicuous, flexibility in the dynamics of social behaviour (for example, patterns of scent marking differ between populations⁹³ of spotted hyaenas^{6,94}, golden jackal³² and otters⁹⁵). Spotted hyaena clans vary between habitats from large, stable matriarchies to small transient associations, and almost every aspect of their social behaviour changes also^{6,94}.

Carnivore societies are the products of varied effects on different individuals of very many selective pressures⁷⁹. There are contemporary carnivore societies, such as those of red and Arctic foxes, badgers, brown hyaenas and farm cats in which stable, well defined relationships exist and yet where the benefits of social ties often appear infrequent or small. These societies share the characteristic that resource dispersion creates conditions in which the smallest range which will support a pair may also support a group. These ecological conditions, which diminish the restraints on group formation, may originally have facilitated the evolution of other types of carnivore society where the contemporary benefits of sociality are conspicuously great, and hence where bigger groups, and thus perhaps ranges larger than the minimum, are advantageous. Species which originally colonized niches in which the ecological constraints upon group formation were lower could expand their niches and hence give rise to new benefits to group living which, in turn, could modify further their social organization. From this viewpoint, the differences in sociality between species are a consequence of interspecific variation in these subsequent benefits of group living. For each society the balance of these contemporary benefits is not necessarily the same as that which originally selected for group living; the origins of the quite different societies of fox and wolf or farm cat and lion may all lie in patterns of prey dispersion which diminished the need to disperse and hence permitted family ties to persist into adulthood and so provided candidates for subsequent and varied selection directed at sociality *per se*. Similarly, the balance of the contemporary benefits to group membership also varies between populations of the same species in accordance with their ecological circumstances and as carnivores are typified by their opportunism, many species have broad niches and hence highly flexible social systems. Elucidating the contemporary selective pressures that promote such social groups requires exploration of the limits to the flexibility of each species' society, and also the processes whereby individuals develop⁹⁶ into quite different social beings depending on the ecological (and social) circumstances into which they are born.

I thank Drs P. J. Bacon, B. C. R. Bertram, T. H. Clutton Brock, S. Erlinge, J. L. Gittleman, J. R. Krebs, H. Kruuk, P. D. Moehlan, J. Rood and D. R. Voigt for stimulating discussions of an earlier version of this paper, and particularly my colleagues in the 'Foxlot' of the Animal Behaviour Research Group at Oxford for useful comments. Many of these ideas on carnivore societies have developed over several years during conversations with Hans Kruuk, without whose guidance and encouragement they would never have crystallized.

1. Rasa, O. A. E. *Z. Tierpsychol.* **43**, 337–406 (1977).
2. Gorman, M. L. *J. Zool. Lond.* **187**, 65–73 (1979).
3. Fox, M. W. *Behaviour of Wolves, Dogs and Related Canids* (Jonathan Cape, London, 1970).
4. Ewer, R. F. *The Carnivores* (Weidenfeld & Nicholson, London, 1973).
5. Kortlandt, A. *Curr. Anthropol.* **6**, 320–326 (1965).
6. Kruuk, H. *Spotted Hyaena, a Study of Predation and Social Behaviour* (University of Chicago Press, Chicago, 1972); in *Function and Evolution in Behaviour* (eds Baerends, G., Beer, C. & Manning, A.) 119–141 (Clarendon, Oxford, 1975).
7. Rood, J. P. *Nature* **248**, 176 (1974); *E. Afr. Wildl. J.* **13**, 89–110 (1978); *Z. Tierpsychol.* **48**, 277–287 (1978).
8. Kruuk, H. *J. Zool. Lond.* **184**, 1–19 (1978); *Behav. Ecol. Sociobiol.* **4**, 75–89 (1978).
9. Kruuk, H. & Parish, T. *J. Anim. Ecol.* **50**, 773–788 (1981).
10. Duplaix, N. *Terre Vie* **34**, 496–620 (1980).
11. Macdonald, D. W. in *Proc. Worldwide Furbearer Conf. Vol 2* (eds Chapman, J. & Pursley, D.) 918–949 (University of Maryland Press, 1981).
12. Macdonald, D. W. & Apps, P. *J. Carnivore Genet. Newsl.* **7**, 256–269 (1978).
13. Dards, J. L. *Carnivore Genet. Newsl.* **7**, 242–255 (1978); thesis, Univ. Southampton (1980).
14. Charles-Dominique, P. *Terre Vie* **32**, 477–528 (1978).
15. Mills, M. G. L. *Carnivore* **1**, 1–7 (1978); *Z. Tierpsychol.* **48**, 113–141 (1978); *J. Zool. Lond.* **198**, 39–51 (1982).

16. Lack, D. *Ecological Adaptations to Breeding in Birds* (Chapman & Hall, London, 1968).
17. Wynne-Edwards, V. C. *Animal Dispersion in Relation to Social Behaviour* (Oliver & Boyd, Edinburgh, 1962).
18. Orians, G. H. *Am. Nat.* **103**, 589–603 (1969).
19. Horn, H. S. *Ecology* **49**, 683–692 (1968).
20. Emlen, S. T. & Oring, L. W. *Science* **197**, 215–223 (1977).
21. Bradbury, J. W. & Vehrencamp, S. L. *Behav. Ecol. Sociobiol.* **1**, 383–404 (1976).
22. Clutton Brock, T. J. *Nature* **250**, 539–542 (1974).
23. Clutton Brock, T. J. & Harvey, P. H. *J. Zool. Lond.* **183**, 1–39 (1977).
24. Hladick, C. M. in *Primate Ecology*, 324–353 (ed. Clutton Brock, T. H.) (Academic, London, 1977).
25. Jarman, P. J. *Behaviour* **58**, 215–267 (1974).
26. Kleiman, D. G. & Eisenberg, J. F. *Anim. Behav.* **21**, 637–659 (1973).
27. Bertram, B. C. R. *E. Afr. Wildl. J.* **11**, 215–225 (1973).
28. Macdonald, D. W. thesis, Univ. Oxford (1977).
29. Hersteinsson, P. & Macdonald, D. W. *Symp. zool. Soc. Lond.* **49**, 259–289 (1982).
30. Kruuk, H. & Parish, T. J. *J. Zool. Lond.* **196**, 31–39 (1982).
31. Clutton Brock, T. J. & Harvey, P. H. *Nature* **273**, 191–195 (1978).
32. Macdonald, D. W. *Behav. Ecol. Sociobiol.* **5**, 17–38 (1979).
33. Macdonald, D. W. et al. in *A Handbook on Biotelemetry and Radio Tracking* (eds Amlaner, C. J. & Macdonald, D. W.) 405–424 (Pergamon, Oxford, 1980).
34. Macdonald, D. W. *Z. Tierpsychol.* **52**, 170–201 (1980).
35. Brown, J. L. *Wilson Bull.* **70**, 160–169 (1964).
36. Macdonald, D. W. *Nature* **282**, 69–71 (1979); in *The Red Fox* (ed. Zimen, E.) 123–175 (Junk, The Hague, 1980).
37. Davies, N. B. in *Behavioural Ecology: an Evolutionary Approach* (eds Krebs, J. R. & Davies, N. B.) 317–350 (Blackwell, Oxford, 1978).
38. von Schantz, T. *Oikos* **37**, 63–68 (1981); thesis, Univ. Lund (1981).
39. Lindstrom, E. in *The Red Fox* (ed. Zimen, E.) 177–184 (Junk, The Hague, 1980); thesis, Univ. Stockholm (1982).
40. Storm, G. L. et al. *Wildl. Monogr.* **49**, 1–82 (1976).
41. Packard, J., Packard, J. M. & Mech, L. D. in *Biosocial Mechanisms of Population Regulation* (eds Cohen, M. N., Malpasse, E. & Klein, H. G.) 135–150 (Yale University Press, 1980).
42. Englund, J. *Viltrevy* **8**, 1–82 (1970).
43. Harrington, F. H., Paquet, P. C., Ryon, J. & Fentress, J. C. in *Wolves of the World* (eds Harrington, F. H. & Paquet, P. C.) 209–222 (Noyes, 1982).
44. Moehlman, P. D. *Nature* **277**, 382–383 (1979); in *Spec. Publ. Am. Soc. Mammal.* **7** (eds Eisenberg, J. R. & Kleiman, D. G.) (in the press).
45. Lomnicki, A. *Oikos* **35**, 185–193 (1980).
46. Bailey, T. N. J. *Wildl. Mgmt* **38**, 438–446 (1974).
47. Erlinge, S. *Oikos* **25**, 308–314 (1974); **28**, 32–42 (1977).
48. Magoun, A. *Acta. zool. fenn.* (in the press).
49. Hornocker, M. G. & Hash, H. S. *Can. J. Zool.* **59**, 1286–1301 (1981).
50. Powell, R. A. *Z. Tierpsychol.* **50**, 153–165 (1979); *The Fisher*, 217 (University of Minnesota Press, 1982).
51. Leyhausen, P. *Symp. zool. Soc. Lond.* **14**, 249–263 (1963).
52. Frame, L. H. & Frame, G. W. *Nature* **263**, 227–229 (1976).
53. Jonsingh, A. J. T. thesis, Indian Inst. Science, Bangalore (1980); *J. Zool. Lond.* (in the press).
54. Mech, L. D. *The Wolf* (Natural History Press, New York, 1970).
55. Bowan, W. D. *Can. J. Zool.* **59**, 639–652 (1981); thesis, Univ. British Columbia (1978).
56. Wyman, J. *Animals* **10**, 79–83 (1967).
57. Lamprecht, J. *Z. Tierpsychol.* **45**, 260–289 (1979).
58. Nel, J. A. J. *Bull. Carnegie Mus. nat. Hist.* **6**, 132–137 (1978).
59. Gittleman, J. L. thesis, Univ. Sussex (1982).
60. O'Farrell, T. *Acta zool. fenn.* (in the press).
61. van Orsdol, K. thesis, Univ. Cambridge (1981).
62. Lamprecht, J. *Mammal Rev.* **11**, 169–179 (1981).
63. Eaton, R. L. *Carnivore* **2**, 9–16 (1979).
64. Lamprecht, J. *Z. Tierpsychol.* **46**, 337–343 (1978); *Z. Säugetierk.* **43**, 210–223 (1978).
65. Schaller, G. B. *The Serengeti Lion: A Study of Predator Prey Relations* (University of Chicago Press, Chicago, 1972).
66. Malcolm, J. thesis, Univ. Harvard (1980).
67. Malcolm, J. & Marten, K. *Behav. Ecol. Sociobiol.* **10**, 1–13 (1982).
68. Frame, L. et al. *Z. Tierpsychol.* **50**, 225–249 (1979).
69. Frame, G. & Frame, L. *Swift and Enduring* 242 (Dutton, New York, 1981).
70. Andelt, W. F. (cited in ref. 55).
71. Bekoff, M. & Wells, M. C. *Scient. Am.* **242**, 130–148 (1980).
72. Camenzind, F. J. in *Coyotes: Biology, Behaviour and Management* (ed. Bekoff, M.) 267–294 (Academic, New York, 1978).
73. Bygott, D. et al. *Nature* **282**, 839–841 (1980).
74. Sunquist, M. thesis, Univ. Minnesota (1979).
75. Sargeant, A. B. *J. Wildl. Mgmt* **36**, 225–236 (1972).
76. Kenward, R. E. *J. Anim. Ecol.* **47**, 449–460 (1978).
77. Hamilton, W. D. *J. theor. Biol.* **7**, 1–52 (1964).
78. Macdonald, D. W. & Moehlman, P. D. in *Perspectives in Ethology* Vol. 5 (eds Bateson, P. P. G. & Klopfer, P.) 433–466 (Plenum, New York, in the press).
79. Bertram, B. C. R. in *Growing Points in Ethology* (eds Bateson, P. P. G. & Hinde, R. A.) 281–301 (Cambridge University Press, 1976).
80. Russell, J. K. thesis, Univ. N. Carolina (1979).
81. Packer, C. & Pusey, A. E. *Nature* **296**, 740–742 (1982).
82. Bertram, B. C. R. in *Serengeti Ecosystem* (eds Norton-Griffiths, M. & Sinclair, A.) 222–248 (Chicago University Press, 1980).
83. Ryon, J. & Fentress, J. C. in *Wolves: a Worldwide Perspective of Their Ecology, Behaviour and Conservation* (eds Harrington, F. H. & Paquet, P. L.) (Noyes Publications, New Jersey, 1982).
84. Rasa, O. A. E. *Z. Tierpsychol.* **42**, 337–342 (1976); *Z. Säugetierknd.* **42**, 108–112 (1977).
85. Mills, M. G. L. *Behav. Ecol. Sociobiol.* **10**, 131–136 (1982).
86. van Lawick, H. & van Lawick Goodall, J. *The Innocent Killers* (Collins, London, 1970).
87. Harrington, F. H. & Mech, L. D. in *Anim. Behav. Soc.*, Fort Collins (1980); in *Wolves of the World* (eds Harrington, F. H. & Paquet, P. C.) 81–105 (Noyes, 1982).
88. Trivers, R. L. in *Sexual Selection and the Descent of Man* (ed. Campbell, B. G.) 378 (Heinemann, London, 1972).
89. Axelrod, R. & Hamilton, W. D. *Science* **211**, 1390–1396 (1981).
90. Mech, L. D. *USDA Forest Service Pap.* NC-97, 10 (1973).
91. Cheeseman, C. L. et al. *J. appl. Ecol.* **18**, 795–804 (1981).
92. Messick, J. P. & Hornocker, M. G. *Wildl. Monogr.* **76**, 1–532 (1981).
93. Macdonald, D. W. *Proc. Symp. zool. Soc. Lond.* **45**, 107–139 (1980); in *Mammalian Social Odours* (eds Brown, R. E. & Macdonald, D. W.) (Oxford University Press, Oxford, in the press).
94. Bearder, S. W. & Randall, R. M. *Carnivore* **1**, 32–48 (1978).
95. Kruuk, H. & Hewson, R. *J. Zool. Lond.* **185**, 205–212 (1978).
96. Bateson, P. P. G. in *Behaviour Development* (eds Immelman, K., Barlow, G., Petrionovich, L. & Main, M.) 1–21 (Cambridge University Press, Cambridge, 1980).
97. Gittleman, J. L. & Harvey, P. H. *Behav. Ecol. Sociobiol.* **10**, 57–63 (1982).
98. Zimen, E. *Z. Tierpsychol.* **40**, 300–341 (1976).

ARTICLES

Variations in magnetization intensity and low-temperature titanomagnetite oxidation of ocean floor basalts

Ulrich Bleil

Institut für Geophysik, Ruhr-Universität Bochum, Postfach 10 21 48, 4630 Bochum 1, FRG

Nikolai Petersen

Institut für allgemeine und angewandte Geophysik, Ludwig-Maximilians-Universität München, Theresienstrasse 41, 8000 München 2, FRG

The remanent magnetization of the oceanic crust exhibits a systematic long-term variation which correlates with the amplitudes of marine magnetic anomalies. After a sharp initial decrease of natural remanent magnetization intensity, a minimum is reached at ~20 Myr, followed by a gradual increase up to ages of 120 Myr. The progressive sea floor alteration of the magnetic minerals carrying the crustal magnetism is proposed as a cause for this behaviour.

THE central anomaly over the spreading axis of active oceanic ridges is by far the most prominent in the pattern of marine magnetic anomalies. The flanking anomalies gradually decrease in amplitude for some distance away from the crest¹. Another characteristic of the amplitude variation in linear marine magnetic anomalies seems to have been overlooked or, where noticed, not explained. This is the increase in amplitude of magnetic anomalies older than ~20 Myr (refs 2–6). Worldwide most of the amplitudes show a pronounced minimum over the age range of 10–30 Myr beyond which, on average, they pro-

gressively increase again to another maximum near the end of the Cretaceous magnetic quiet zone (Fig. 1). This tendency seems to be independent of their geographical location or the spreading rate of a ridge.

At the onset of the Mesozoic M-series magnetic anomalies at ~108 Myr the amplitudes are again high in most cases. Longer profiles of the M-series show a steady diminution with age in the amplitudes of the anomalies, leading to a gradual rather than an abrupt transition into the Jurassic magnetic quiet zone^{7,8}.

Magnetization variation with age in oceanic basalts

The intensity of the natural remanent magnetization (NRM) of the upper 500 m of the oceanic crust shows almost identical but even clearer systematic temporal variations to those described above for the marine magnetic anomalies. A compilation of NRM data for drill cores of ocean crust recovered by the Deep Sea Drilling Project is shown in Fig. 2. It summarizes all results available to us up to DSDP Leg 65 and comprises 126 geometric site mean values based on individual measurements of almost 3,500 samples. The magnetic intensities given are reduced to the Equator using basement ages inferred from biostratigraphical dating of overlying sediments, magnetic anomaly patterns, or both, together with global plate reconstruction models. At drill sites where less than 10 samples were measured, the NRM data were combined with those of the nearest neighbouring site to yield a more reliable mean. Details of this summary are given elsewhere⁹.

After a sharp initial drop in NRM intensity during the first 3–5 Myr, the magnetization of the upper oceanic crust passes through a minimum between ~10 and 30 Myr. In contrast, there is a continuous increase of the average NRM values over the entire age range from 30 to 120 Myr followed once more by a decrease in Jurassic time. The latter decrease, however, is documented at present by only a few data points. Comparison of Figs 1 and 2 shows that the trends seen in both the magnetic anomaly amplitudes and the crustal magnetization intensity correlate not only on a global scale but also for all the major ocean basins over the period of time represented.

This coincidence of anomaly-amplitude variation and ocean-crust-magnetization variation leads to the important conclusion that the main sources for the oceanic magnetic anomaly lineations are represented by the types of basement rocks recovered by the DSDP, that is predominantly extrusive basaltic rocks that are affected by steady low-temperature sea floor alteration of their magnetic minerals over time. This contrasts with the view that deep seated intrusive rocks contribute significantly to the anomalies^{10–14} but supports the statistical emplacement model of the magnetic source layer¹⁵. Furthermore, the Cretaceous magnetic quiet zone is obviously not caused by weakly magnetized upper oceanic crust, but rather by a long period of uniform polarity of the Earth's magnetic field. For the Jurassic magnetic zone the present data are consistent with its originating from a progressive breakdown of the original NRM intensities at this stage.

Causes of magnetization variations of the oceanic crust

Various explanations have been proposed for the anomalous intensity of the central anomaly: (1) gross variations in the amounts of magnetic minerals contained in ocean floor

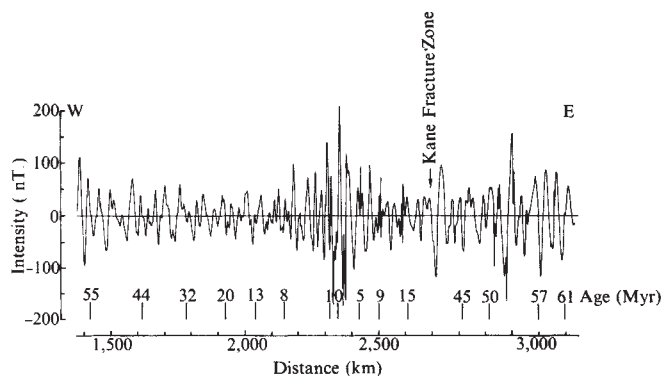


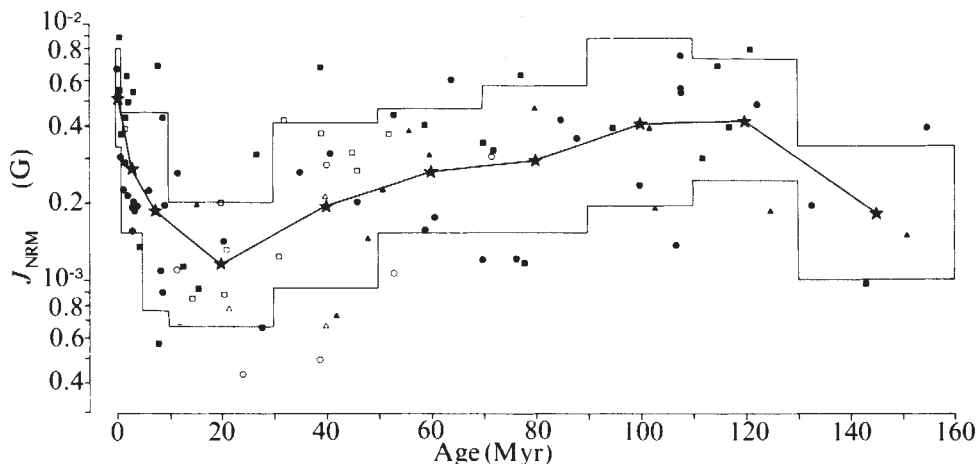
Fig. 1 Magnetic profile across the North Atlantic between Puerto Rico and Canary Islands (solid line in insert) obtained by *Glomar Challenger* during DSDP Leg 46.

basalts¹⁶; (2) spontaneous decay of natural remanent magnetization intensity with time due to thermal agitation^{17,18}; (3) variations in the past strength of the Earth's magnetic field¹⁹; (4) off-axis dyke injection at the ridge crest and consequent contamination of the older reversely magnetized crust on either side of the normally magnetized central block with oppositely magnetized material thereby reducing the bulk intensity of their magnetization²⁰; (5) NRM decay due to progressive low-temperature oxidation of the magnetic minerals contained in the ageing ocean floor basalts^{16,21–23}.

Based on Fig. 2, we can make the following comments:

- (1) The first effect can be dismissed readily. There are no gross variations with distance from the ridge axis where basaltic rocks have been emplaced in the amounts of magnetic minerals^{24–26}.
- (2) Viscous decay of magnetization may contribute to the observed NRM decrease over the first 20 Myr but it cannot account for the subsequent increase of magnetic intensity.
- (3) There is no evidence from palaeointensity determinations that the strength of the Earth's magnetic field has gradually changed by a factor of 10 or so within the past 100 Myr (ref. 27). The strongest evidence against this argument comes from the work of Prévot and Grommé²¹ who compared the NRM intensities of subaerial basalts and ocean floor basalts ranging from historical time to 10 Myr. They found that no important decrease in the intensity of magnetization has taken place in subaerial basalts during this time. Consequently they argue that the diminution of NRM observed in ocean floor basalts is caused by another independent effect—low-temperature oxidation of the titanomagnetites contained in them.
- (4) A statistical distribution of dyke injections about the centre of an oceanic ridge results in contamination of the uniformly magnetized blocks flanking the central block with new material of opposite polarity and, therefore, in principle, can explain

Fig. 2 Mean values (geometric averages reduced to the Equator) for the natural remanent magnetization intensity, J_{NRM} , of ocean floor basalts recovered at DSDP drill sites versus crustal age. ●, North Atlantic (including the Caribbean and Mediterranean Sea); ○, South Atlantic; ▲, Indian Ocean; △, Antarctic Ocean; ■, Pacific Ocean (including the Bering Sea); □, Philippine Sea. The stars denote geometric mean values based on site averages for the time periods indicated by the standard deviation lines.



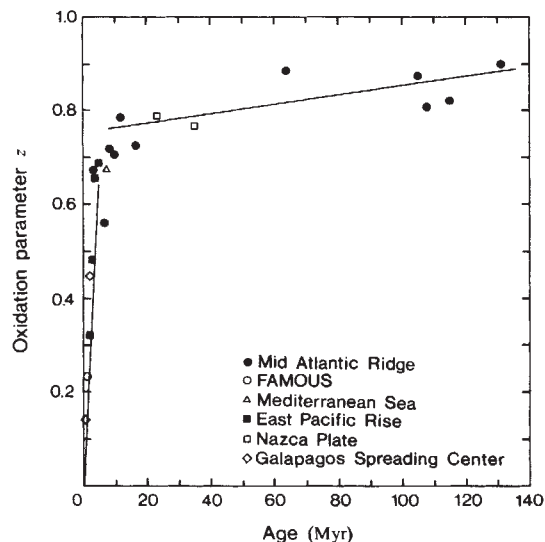


Fig. 3 Mean values of titanomagnetite oxidation parameter z (ref. 41) versus age of ocean floor basalts in various DSDP drill sites and the FAMOUS area.

variations in the amplitudes of the magnetic anomalies. However, this process cannot explain the corresponding variations of magnetization intensity of the individual ocean floor samples shown in Fig. 2 and consequently must also be dismissed as a likely cause. Another argument against the statistical injection of dykes as the main cause for the amplitude variation has been given by Harrison¹⁹. He points out that if the actual dates for field reversals are used, it does not seem possible to choose a plausible areal extent for the dyke injection that will at the same time preserve the anomalies associated with the Jaramillo and Olduvai events and make the central anomaly sufficiently large relative to the surrounding anomalies.

(5) It has been demonstrated that the low-temperature oxidation of magnetic minerals contained in ocean floor basalts is a process that has proceeded steadily with time in all oceanic basalts recovered so far; the oxidation state being higher in older rocks^{23,25,28-31}. According to work on synthetic titanomagnetites, low-temperature oxidation results in a decrease of NRM intensities of ocean floor basalts. The experiments, however, give only very limited evidence of a subsequent magnetization increase for higher oxidation states. Inversion of the metastable cation-deficient titanomagnetite to an intergrowth of magnetite and other stable Fe-Ti oxides could, however, produce an increase in magnetization³⁵.

Low-temperature oxidation of titanomagnetites in ocean floor basalts

Of the mechanisms discussed above, low-temperature oxidation of the magnetic mineral assemblage seems to be the dominant process. We have, therefore, developed a model of low-temperature titanomagnetite oxidation that explains, at least qualitatively, the observed intensity variations and, correspondingly, the variations in amplitude of the marine magnetic anomalies.

At the ridge crest, the most important magnetic mineral in oceanic basalts is a stoichiometric titanium-rich titanomagnetite with some impurities, mainly Mg and Al. Other magnetic minerals, such as pyrrhotite, have only a negligible role in fresh ocean floor basalts. The composition of this primary titanomagnetite varies only within narrow limits and independent of the ridge axis at which they were emplaced^{28,29,36-38}. It comes close to the composition²⁵: $\text{Fe}_{2.27}\text{Ti}_{0.58}\text{Al}_{0.07}\text{Mg}_{0.06}\text{Mn}_{0.02}\text{O}_4$. In sea floor conditions, the primary titanomagnetites are subject to a steady low-temperature oxidation which results in the formation of a cation-deficient spinel phase. The cation-deficiency, on the average, increases with age of the rock.

Figure 3 suggests that the ocean floor titanomagnetite oxidation is not a simple one-stage process as proposed by Ozima

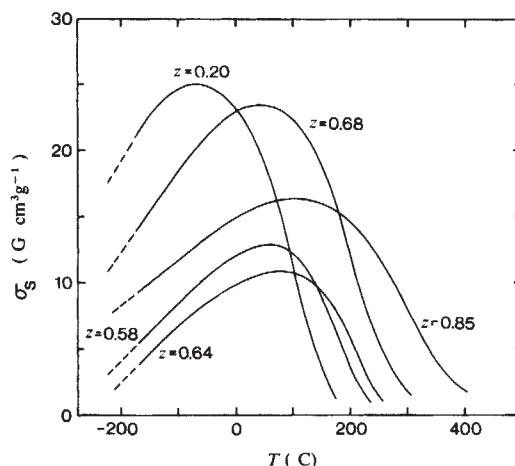
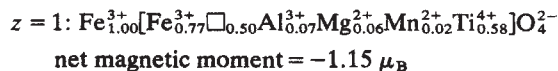
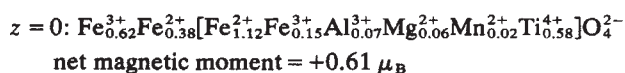


Fig. 4 Magnetic mineral saturation magnetization, σ_s , as function of temperature, T , for titanomagnetites of variable oxidation state extracted from oceanic basalts recovered at DSDP drill sites in the Atlantic.

*et al.*³¹. Instead we observe two approximately linear sections with a distinct change of slope at around 10 Myr. It is likely, therefore, that two different processes oxidize titanomagnetite in suboceanic conditions. If the relationship of Fig. 3 holds for the ocean crust in general, almost two-thirds of the titanomagnetite oxidation must take place within the first 10 Myr and the rest at a much slower rate over a long period of time.

The mechanism underlying the sea floor oxidation of titanomagnetite is Fe-migration out of the spinel lattice^{25,38,39}. Titanium is not affected by this migration process. Contrary to the oxidation model of Readman and O'Reilly³³ this leads to a constantly changing Fe/Ti ratio in the spinel lattice with oxidation. The egressing Fe ions may be either incorporated in surrounding clay minerals or transported elsewhere by seawater⁴⁰. Taking into account this process of Fe-migration during titanomagnetite oxidation, we can calculate the resultant variation of magnetization using the following idealized models:

(A) For the case where vacancies are formed only at octahedral sites of the spinel lattice, there is a decrease of the net magnetic moment at low-oxidation stages, but eventually the spontaneous magnetization reverses polarity and then grows in the opposite direction. This case is tested for the above given primary titanomagnetite composition in oceanic tholeiites assuming a cation distribution according to the O'Reilly-Banerjee model⁴¹. We also make the simplifying assumption that Ti, Al, Mg, Mn and cation vacancies are located only at octahedral sites.



The brackets denote the cations in octahedral sites. The cations outside the brackets are in tetrahedral sites. A cation vacancy in either the octahedral or tetrahedral sublattice is indicated by $\square$. The magnetic moments of the cations in octahedral sites are aligned antiparallel to the magnetic moments of the cations in tetrahedral sites. The net magnetic moment of the spinel structure is therefore calculated from the vector sum of both sublattices, Fe^{3+} having a magnetic moment of 5 Bohr magnetons (μ_B) and Fe^{2+} and Mn^{2+} having 4 and 5 Bohr magnetons, respectively. The other cations, Mg^{2+} and Al^{3+} , do not contribute magnetically. The reversal occurs at $z = 0.60$. The calculated magnetic moment is correct only at 0 K; at room temperature the situation may be different depending on the temperature variation of magnetization (on the Néel type curve⁴²). For model (A) we would expect a sequence of Q-P-L-N-Q'

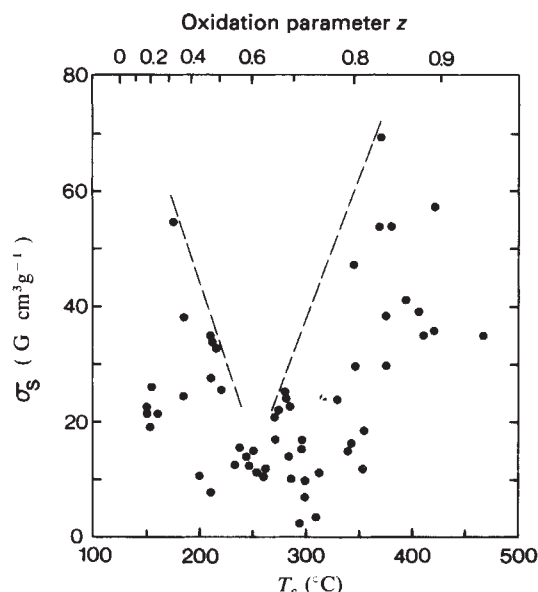
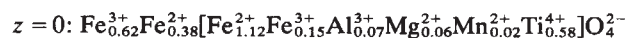


Fig. 5 Magnetic mineral saturation magnetization, σ_s (at room temperature), versus Curie temperature, T_c , and oxidation parameter z , respectively, for titanomagnetites extracted from ocean floor basalts recovered at DSDP drill sites in the Atlantic and Pacific oceans.

Néel curves with increasing oxidation^{32,33,43}. This prediction was tested for several ocean floor basalts (Fig. 4). These results show that the expected sequence is not observed, instead the measured curves suggest a sequence P-L-P.

(B) For the case where vacancies are formed on octahedral sites until $z = 0.60$ is reached (corresponding, approximately, to the change in slope in Fig. 3) after which z -value vacancies are formed on tetrahedral sites only the variations are:



$$\text{net magnetic moment} = +0.61 \mu_B$$

$$z = 0.60:$$



$$\text{net magnetic moment} = +0.01 \mu_B$$



$$\text{net magnetic moment} = +0.86 \mu_B$$

The formation of vacancies at tetrahedral sites of the spinel lattice has been shown experimentally for the oxidation of pure magnetite^{44,45}. It is not yet clear, however, if and in what conditions tetrahedral vacancies are formed in titanomagnetites.

Like model (A), model (B) produces a decrease of magnetization until $z = 0.60$ is reached. For higher z -values, however, before the magnetization reverses, the magnetic moment increases. For such a model we expect a sequence of P-L-P Néel curves which, in fact, are observed (Fig. 4). O'Donovan and O'Reilly³⁴ obtained a similar sequence of Néel curves from progressively oxidized synthetic titanomagnetite of composition $\text{Fe}_{2.35}\text{Mg}_{0.05}\text{Ti}_{0.60}\text{O}_4$. The reported P-type behaviour for this composition becomes more pronounced with increasing z until it is almost L-type at $z = 0.59$. Beyond this z -value, the curves show less pronounced P-type behaviour again. This agrees well with our observations.

Comparison of models (A) and (B) suggests that in conditions at the ocean floor, model (B) is more realistic for the following reasons. First, in model (B) Fe-diffusion from octahedral sites, and then, for z greater than 0.60, from tetrahedral sites accounts for the two-stage oxidation shown by Fig. 3. It explains, at least qualitatively, the different rates of oxidation, as the cations at octahedral sites are more mobile than those at tetrahedral sites⁴⁶. Second, there is no evidence for the self-reversal of

magnetization in ocean floor basalts that should be expected from model (A) when the oxidation parameter z exceeds 0.60 (refs 25, 47). Third, variations of Néel-type curves as discussed above.

In the variation of saturation magnetization of titanomagnetites in ocean floor basalt with increasing oxidation state (Fig. 5) a distinct minimum is seen at about $z = 0.60$. This, however, does not favour either model because in both cases, the amount of magnetization reaches a minimum at $z = 0.60$ and increases again for higher z -values.

Inversion of highly cation-deficient titanomagnetite may contribute to the observed increase in magnetization but the effect does not seem to be the essential one. All the samples plotted in Fig. 5 are single phase cation-deficient spinels, judging by the shape of their thermomagnetic curves (for example, Fig. 4) and their X-ray diffraction patterns²⁵.

Conclusions

The actual mechanism of cation diffusion may be much more complicated than our simple 'model B'. Nevertheless, it shows that a physically reasonable model yielding all the main characteristics of the variations in magnetization observed for the progressive low-temperature oxidation of titanomagnetites in ocean floor basalts is feasible. This type of magnetic mineral alteration appears to proceed steadily with increasing age of the oceanic crust in sea floor conditions. The rate of oxidation, is not constant over time, however, but is fast for about the first 10 Myr and is much slower beyond this age. A direct consequence of the low-temperature oxidation of titanomagnetites in oceanic tholeiites is a corresponding temporal variation in the NRM of the upper, predominantly effusive, basaltic oceanic basement. Such a variation accounts not only for the initial decrease in NRM intensity during the first 10–20 Myr, as found by previous workers, but also explains the subsequent slow but steady increase up to ages of ~120 Myr. The decrease in remanent intensity suggested for the Jurassic is as yet poorly documented, but appears to be beyond the validity range of our oxidation model. If real, this decrease probably results from a disintegration of the highly cation-deficient and thus unstable titanomagnetites into an assemblage of less magnetic minerals.

The temporal variation pattern seen in the magnetization intensity of the upper oceanic basement rocks is apparent also in the general intensity of the linear marine magnetic anomalies, suggesting that both of these phenomena result from the progressive alteration of the titanomagnetites carrying the remanent crustal magnetism. This result confirms previous studies which have shown that the predominant anomaly source layer is confined to the uppermost portion of the oceanic crust—frequently referred to as Layer 2A—and that deeper structures do not effectively contribute to the anomaly amplitudes.

This work was supported by the Deutsche Forschungsgemeinschaft. G. Ross Heath and Lloyd Burckle read the manuscript and made helpful suggestions.

Received 16 August; accepted 19 November 1982.

- Vine, F. J. *Science* **154**, 1405–1415 (1966).
- Heirtzler, J. R. & Le Pichon, X. *J. geophys. Res.* **70**, 4013–4033 (1965).
- Vogt, P. R. & Johnson, G. L. *Nature* **245**, 373–375 (1973).
- Heirtzler, J. R., Dickson, G. O., Herron, E. M., Pitman, W. C. & Le Pichon, X. *J. geophys. Res.* **73**, 2119–2136 (1968).
- Le Pichon, X. & Heirtzler, J. R. *J. geophys. Res.* **73**, 2101–2117 (1968).
- Dickson, G. O., Pitman, W. C. & Heirtzler, J. R. *J. geophys. Res.* **73**, 2087–2100 (1968).
- Hayes, D. E. & Rabinowitz, P. D. *Earth planet. Sci. Lett.* **28**, 105–115 (1975).
- Larson, R. L. & Hilde, T. W. C. *J. geophys. Res.* **80**, 2586–2594 (1975).
- Bleil, U. thesis, Ruhr-Univ. Bochum (1981).
- Ade-Hall, J. M. et al. *Nature* **255**, 389–390 (1975).
- Blakely, R. J. *Am. geophys. Un. Monogr.* **19**, 227–235 (1976).
- Kidd, R. G. W. *Earth planet. Sci. Lett.* **33**, 310–320 (1976).
- Harrison, C. G. A. *Geophys. J. R. astr. Soc.* **47**, 257–284 (1976).
- Angenheister, G., Petersen, N. & Schönharting, G. *Fortschr. Miner.* **54**, 54–92 (1977).
- Schouten, H. & Denham, C. R. *Am. geophys. Un., Maurice Ewing Ser.* **2**, 151–159 (1979).
- Irving, E. *Can. J. Earth Sci.* **7**, 1528–1538 (1970).
- Irving, E., Robertson, W. A. & Aumento, F. *Can. J. Earth Sci.* **7**, 226–238 (1970).
- Carmichael, C. M. *Can. J. Earth Sci.* **7**, 239–256 (1970).
- Harrison, C. G. A. *J. geophys. Res.* **73**, 2137–2142 (1968).
- Matthews, D. H. & Bath, J. *Geophys. J. R. astr. Soc.* **13**, 349–357 (1967).
- Prévot, M. & Grommé, S. *Geophys. J. R. astr. Soc.* **40**, 207–224 (1975).
- Johnson, H. P. & Atwater, T. *Bull. geol. Soc. Am.* **88**, 637–647 (1977).

23. Prévot, M., Lecaille, A. & Hekinian, R. *Am. geophys. Un. Maurice Ewing Ser. 2*, 210–229 (1979).
24. Haggerty, S. E. *Trans. Am. geophys. Un.* **51**, 273–277 (1970).
25. Petersen, N., Eisenach, P. & Bleil, U. *Am. geophys. Un. Maurice Ewing Ser. 2*, 169–209 (1979).
26. Eisenach, P. thesis, Ludwig-Maximilians-Univ. München (1980).
27. Smith, P. J. *Earth planet. Sci. Lett.* **2**, 329–330 (1967).
28. Bleil, U. & Petersen, N. *Init. Rep. DSDP, Leg 37*, 449–456 (1977).
29. Johnson, H. P. & Hall, J. M. *Geophys. J. R. astr. Soc.* **52**, 45–64 (1978).
30. Schaeffer, R. M. & Schwarz, E. J. *Can. J. Earth Sci.* **7**, 268–273 (1970).
31. Ozima, M., Joshima, M. & Kinoshita, H. *J. Geomagn. Geoelectr.* **26**, 335–354 (1974).
32. Ozima, M. & Sakamoto, N. *J. geophys. Res.* **67**, 7035–7046 (1971).
33. Readman, P. W. & O'Reilly, W. J. *Geomagn. Geoelectr.* **24**, 69–90 (1972).
34. O'Donovan, J. B. & O'Reilly, W. *Earth planet. Sci. Lett.* **34**, 291–299 (1977).
35. Readman, P. W. & O'Reilly, W. *Phys. Earth planet. Inter.* **4**, 121–128 (1970).
36. Prévot, M. & Lecaille, A. *Bull. Soc. geol. Fr.* **19**, 903–911 (1976).
37. Ade-Hall, J. M., Johnson, H. P. & Ryall, P. J. C. *Init. Rep. DSDP Leg 34*, 459–468 (1976).
38. Johnson, H. P. & Melson, W. G. *Init. Rep. DSDP Leg 45*, 575–579 (1978).
39. Prévot, M., Remond, G. & Caye, R. *Bull. Soc. Fr. Miner. Crist.* **91**, 65–74 (1968).
40. Böhlke, J. K., Honnorez, J., Honnorez-Guerstein, B. M., Muehlenbachs, K. & Petersen, N. *J. geophys. Res.* **86**, 7935–7950 (1981).
41. O'Reilly, W. & Banerjee, S. K. *Miner. Mag.* **36**, 29–37 (1967).
42. Néel, L. *Ann. Phys.* **3**, 137–198 (1948).
43. Schult, A. *Earth planet. Sci. Lett.* **4**, 57–63 (1968).
44. Kullerød, G. & Donnay, G. Z. *Kristallogr.* **128**, 1–17 (1969).
45. Weber, H. P. & Hafner, S. Z. *Kristallogr.* **133**, 327–340 (1971).
46. Fröhlich, F. & Stiller, H. *Geofis. Pur. Appl.* **55**, 91–100 (1963).
47. Hall, J. M. *J. Geomagn. Geoelectr.* **29**, 411–420 (1979).

Expression of complete transplantation antigens by mammalian cells transformed with truncated class I genes

Robert S. Goodenow*, Iwona Stroynowski*, Minnie McMillan[†], Margery Nicolson[‡], Kurt Eakle*, Beverly Taylor Sher*, Norman Davidson[§] & Leroy Hood*

* Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

[†] Department of Microbiology, USC Medical School, Los Angeles, California 90033, USA

[‡] AMGen, Newbury Park, California 91320, USA

[§] Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA

Mouse L cells transformed with the cloned class I genes of the major histocompatibility complex of the mouse express transplantation antigens with serological determinants of the donor haplotype. However, transformation with the truncated subclones of a BALB/c H-2L^d gene containing the exons encoding the external domains also leads to the production of cells which express complete cell-surface molecules. Moreover, full-length products of the foreign haplotype, as judged by serological and biochemical criteria, are generated independently of the use of carrier DNA in transformation. However, the frequency of productive transformation is substantially less than that obtained with a complete gene. The most plausible explanation for these phenomena involves homologous recombination between host chromosomal and donor class I sequences.

THE class I genes of the major histocompatibility or H-2 complex of the mouse encode a family of polymorphic cell-surface molecules known as the transplantation antigens. These antigens, first defined on the basis of allograft rejection^{1,2}, are involved in the T-cell recognition of cells altered by viral infection or neoplastic transformation (for reviews see refs 3, 4). The transplantation antigens are highly polymorphic. As different strains of mice have distinct combinations of class I alleles (haplotypes), specific alloantisera and monoclonal antibodies have been prepared by the appropriate cross-immunizations with cells from inbred, congenic, or recombinant strains of mice. The serological analyses have demonstrated that each strain contains particular class I molecules which can be indicated by the appropriate capital letters with small superscript letter to designate haplotype. For example, the BALB/c mouse, which is of the H-2^d haplotype, has at least three major transplantation antigens denoted K^d, D^d and L^d.

The class I molecules are glycoproteins of molecular weight (MW) 45,000 which can be expressed on the cell surface only when associated noncovalently with the 12,000-MW polypeptide β_2 -microglobulin^{5,6}. The class I polypeptide can be divided on the basis of its structure into three external domains, each of about 90 amino acids, a transmembrane region of approximately 40 residues, and one cytoplasmic domain of 30 residues^{7,8}. The first and second external domains appear to be the most variable among different class I molecules and probably contain the highly polymorphic determinants defined serologically. The third external domain is the most highly conserved among the different class I molecules⁹ and is associated with β_2 -microglobulin. Although the transmembrane and cytoplasmic domains are less conserved than the third domain, there are no serologically defined determinants which can dis-

tinguish the C-terminal regions of different class I molecules or molecules of different haplotypes.

The isolation of cDNA probes specific for the class I genes¹⁰ and the characterization of genomic clones obtained using these probes revealed that the domain organization of a class I polypeptide is reflected in the exon-intron structure of the class I gene¹¹. The first exon encodes the leader sequence, while the second, third and fourth exons encode, respectively, the first, second and third external domains (Fig. 1). The transmembrane region is encoded by the fifth exon and the small cytoplasmic domain by the sixth, seventh and eighth exons.

The coding assignments for a number of cloned class I genes isolated from BALB/c lambda (λ) and cosmid genomic libraries⁹ have been determined using gene transfer^{12,13}. For example, the H-2L^d gene was identified through the production of L^d molecules by thymidine kinase negative (tk⁻) mouse L cells (H-2^k haplotype) cotransformed with the herpes viral (HSV) tk gene and DNA from the BALB/c λ clone 27.5 (ref. 12). The H-2^k haplotype class I gene products of mouse L cells can readily be distinguished from their BALB/c counterparts of the H-2^d haplotype by appropriate alloantisera and monoclonal antibodies. The nucleic acid sequence of the 27.5 gene has been determined and encodes a translated amino acid sequence identical to that of the L^d antigen at all 77 identified residues¹⁴. Gene transfer has also been successful in identifying the BALB/c genes encoding other serologically defined class I antigens, such as K^d and D^d, as well as the related genes of the Tla complex encoding the haematopoietic differentiation antigens TL and Qa¹³. Similarly, other laboratories have used gene transfer to identify specific cloned class I genes from other substrains¹⁵ or strains¹⁶ of mice.

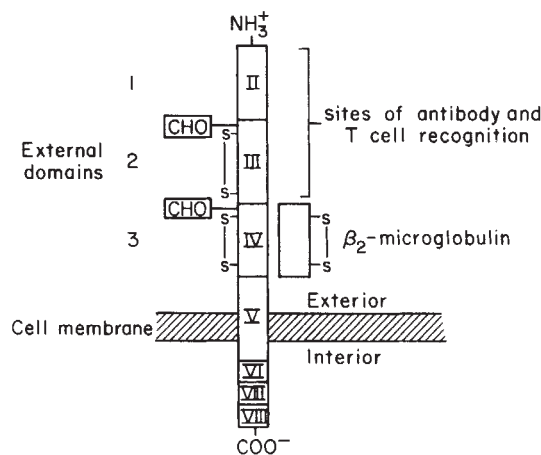


Fig. 1 Schematic representation of the structural organization of class I molecules. The class I molecule is represented as a vertical bar, transverse the cell membrane which is divided into the exons (indicated as Roman numerals) encoding each of the domains numbered at the left. The first exon encoding the leader peptide is not shown. The subunit β_2 -microglobulin is represented as a small rectangle at the lower right. The points of attachment of carbohydrate are indicated at the left of the diagram as boxes labelled CHO. Disulphide bridges are marked with the letter 'S'.

Table 1 Clonal analysis of H-2L^d expression by λ 27.5 cotransformants

Clone designation	c.p.m.*
K2a.1	18.9
.2	15.3
.3	17.9
.4	18.1
.5	18.8
.6	19.1
.7	20.1
.8	17.5
.9	20.0
.10	18.0
.11	18.8
.12	19.0
.13	18.9
.14	19.0
.15	17.8
Ltk ⁺ (pooled)	$\bar{N} = 18.48 \pm 0.25$

Approximately 10^7 thymidine kinase negative (tk⁻) mouse L(Ltk⁻) cells, maintained in α -MEM medium (Irvine) containing 10% fetal calf serum (FCS), glutamine, antibiotics and $300 \mu\text{g ml}^{-1}$ bromodeoxyuridine, were transformed as previously described¹² with 20 ng of the HSV tk gene in pBR322 (ptK5) and $1.0 \mu\text{g}$ 27.5 DNA with $20 \mu\text{g}$ of carrier DNA. The tk positive (tk⁺) transformants were selected in medium containing hypoxanthine, aminopterin and thymidine (HAT)³⁰ as described previously¹². Individual clones of the various transformants were obtained by cloning in soft agar. Approximately 10^3 cells were seeded into an equal volume (0.25 ml) of 0.5% low temperature melting agar (BioRad) containing 20% FCS in HAT medium. The soft agar was pre-chilled to 45°C before the addition of cells and immediately plated into Petri dishes. Single colonies were randomly picked with microcapillary pipettes and transferred to microtitre plates before passage in mass culture. Clones were selected from cultures which had been passaged less than two times following trypsin dispersion of the primary colonies to avoid selective loss of rare cotransformants. Transformants were trypsinized and assayed with monoclonal hybridoma antibodies as previously described¹² except for the addition of 150–200,000 c.p.m. of ^{125}I -protein A (Amersham). Ltk⁺ cells are tk transformants obtained without cloned class I DNA.

* Average numbers of c.p.m. ($\times 10^{-3}$) of ^{125}I -protein A bound per 5×10^5 cells with a 10^{-2} dilution of the 30–57 monoclonal hybridoma antibodies to L^d (ref. 31). Counts per min vary from the calculated mean by no more than 3% of the input c.p.m. (160,000).

We describe here the transformation of L cells using sub-cloned fragments of a BALB/c class I gene which result in the synthesis of complete class I polypeptides of the donor haplotype. These results can best be explained by homologous recombination between the transforming gene fragments and host chromosomal class I genes.

Transformation with the complete H-2L^d gene

To study transformation with truncated class I genes, we first analysed the various parameters involved in transformation with complete genes. Cotransformation of 10^7 Ltk⁻ cells with HSV tk DNA ($0.02 \mu\text{g}$), Ltk⁻ carrier DNA ($20 \mu\text{g}$) and $1.0 \mu\text{g}$ of λ 27.5 DNA, which contains the entire BALB/c H-2L^d gene, results in the production of tk positive (tk⁺) transformant colonies which are almost all L^d positive (Table 1). In addition, the expression of L^d molecules by tk⁺ cotransformants has also been analysed by fluorescence staining with monoclonal antibodies and quantitation on the fluorescence-activated cell sorter. More than 90% of the tk⁺ cotransformants express similar levels of L^d molecules whereas 10% of these transformants express significantly higher or lower numbers of these gene products (data not included).

The effects of varying the amount of transforming DNA used on the frequency of L^d positive transformation were determined as follows. Ltk⁻ cells were cotransformed with varying amounts of λ 27.5 DNA and a constant amount of HSV tk DNA and homologous Ltk⁻ carrier DNA as described previously¹³. The resulting tk⁺ transformants were analysed for L^d expression in the cell binding radioimmunoassay¹³. Figure 2a shows that L^d cotransformation, as measured by the magnitude of the radioimmunoassay signal, is dependent on the concentration donor DNA.

To determine whether the numbers of L^d cotransformants in uncloned tk⁺ cultures could be estimated by radioimmunoassay we did the following experiments. A cloned L^d transformant line, designated 2a-7 (see Fig. 2b legend), was added at different ratios to tk⁺, L^d negative transformants (obtained from transformation with the tk gene alone) and the reconstituted samples assayed by radioimmunoassay for L^d expression. Figure 2b shows that the magnitude of the radioimmunoassay signal is dependent on the number of L^d positive cells in the sample. Moreover, the limit of sensitivity of this assay for detecting positive cells is about 0.1%. To confirm the estimated per cent L^d positive cotransformants, three cell lines derived from transformation with 1.0, 0.1 and $0.01 \mu\text{g}$ of λ 27.5 DNA (Fig. 2a), were cloned and tested individually in the radioimmunoassay. Table 2 shows that the number of L^d cotransformants and the

Table 2 Clonal analysis of the frequency of cotransformation versus the amount of transforming DNA

1.0 μg λ 27.5		0.1 μg λ 27.5		0.01 μg λ 27.5	
Clone	c.p.m.	Clone	c.p.m.	Clone	c.p.m.
12-10.1	17.2	12-11.1	0.2	12-12.1	0.2
.2	14.9	.2	10.9	.2	0.3
.3	6.7	.3	15.2	.3	0.2
.4	14.6	.4	0.3	.4	0.2
.5	15.9	.5	15.4	.5	0.2
.6	14.7	.6	6.2	.6	0.3
.7	16.1	.7	13.9	.7	0.2
.8	15.2	.8	15.6	.8	0.4
.9	14.7	.9	0.2	.9	0.2
.10	16.0	.10	0.2	.10	0.2

The L^d transformants generated in the experiment depicted in Fig. 2a were cloned as described in the legend to Table 1 and tested in the radioimmunoassay with monoclonal antibodies to L^d molecules. The c.p.m. indicate the average numbers of c.p.m. ($\times 10^{-3}$) ^{125}I -protein A bound per 5×10^5 cells with a 10^{-2} dilution of 30–57 hybridoma ascites fluid. The per cent L^d positive clones are 100%, $1.0 \mu\text{g}$; 60%, $0.1 \mu\text{g}$; and less than 10% (that is, fewer than 1 in 10), $0.01 \mu\text{g}$.

percent L^d positive cells in a pooled culture estimated from the standard curve (Fig. 2b) agrees with the number of L^d positive clones obtained from such mixed populations. Surprisingly, the positivity of the individual L^d cotransformants, as estimated from the magnitude of the radioimmunoassay signal, was independent of the amount of $\lambda 27.5$ DNA used.

The effects of using homologous Ltk^- DNA as carrier during transformation were also investigated. When carrier DNA was omitted from cotransformations, the numbers of tk^+ colonies dropped dramatically depending on the amount of $\lambda 27.5$ DNA used (Fig. 2c). The per cent L^d positive cells obtained per μg of $\lambda 27.5$ DNA and number of tk^+ colonies paralleled the results obtained for transformation with carrier (Fig. 2c). The use of higher $\lambda 27.5$ DNA concentrations without carrier partially restores the production of tk^+ colonies indicating that at high concentrations the transforming DNA could serve as carrier for the purposes of colony formation (Fig. 2c). Cotransformation with a constant amount of HSV tk DNA and vector sequences alone (for example, pBR322 or λ DNA) at high concentrations also reduced the frequency of tk^+ transformation (data not included). These effects were not studied further.

In summary, these results indicate that the number of L^d positive transformants increases to a plateau value with increasing concentrations of transforming L^d DNA. At high concentrations of L^d DNA (that is, $>0.1 \mu g$), the frequency of L^d to tk^+ cotransformation is at least 90%. The level of L^d expression per cotransformed cell appears to be essentially independent of the amount of L^d DNA used. The reasons for this are unknown. When carrier DNA is omitted, the frequencies of L^d and of tk^+ transformation are both reduced, although the relative ratio between L^d and tk^+ transformation remains high.

Transformation with a truncated $H-2$ gene

In the course of screening the BALB/c library for the genes encoding the serologically defined class I molecules¹³, by transformation, a truncated class I gene which yielded K^d positive cotransformants was identified. One out of 100 of the tk^+ transformants produced in the gene transfer with the HSV tk gene and $\lambda 19.2$ DNA reacted with the monoclonal antibodies to K^d molecules (data not shown). These results were surprising because restriction map analyses demonstrated that the 19.2 clone contained no more than the first three exons of a class I gene together with 5' flanking sequences. Indeed, no hybridization was observed between $\lambda 19.2$ DNA and a class I cDNA probe specific for exons 4 to 8 (Fig. 3A). Therefore, we conclude that all five of the 3' most exons of a class I gene are not present on the $\lambda 19.2$ insert. Since the $\lambda 19.2$ phage was purified from a single plaque at least twice before use, it is unlikely that clone 19.2 is contaminated with DNA containing a complete class I gene. In addition, the sensitivity of hybridization with the radio-labelled cDNA probe used for restriction mapping should permit detection of as few as 10 copies of class I clone DNA. Thus, the contamination of 19.2 DNA with less than 0.0001% of another class I gene would have been detected by restriction mapping and be insufficient for positive transformation.

Two-dimensional gel electrophoretic analyses of the products of the gene $\lambda 19.2$ were striking in two regards (Fig. 3B). First, the class I polypeptide reacting with the monoclonal K^d antibodies had a molecular weight of $\sim 45,000$, the size of an intact class I polypeptide. Second, the charge distribution of these class I polypeptides was almost identical to that of the normal K^d polypeptides. These results indicate that the L cell can

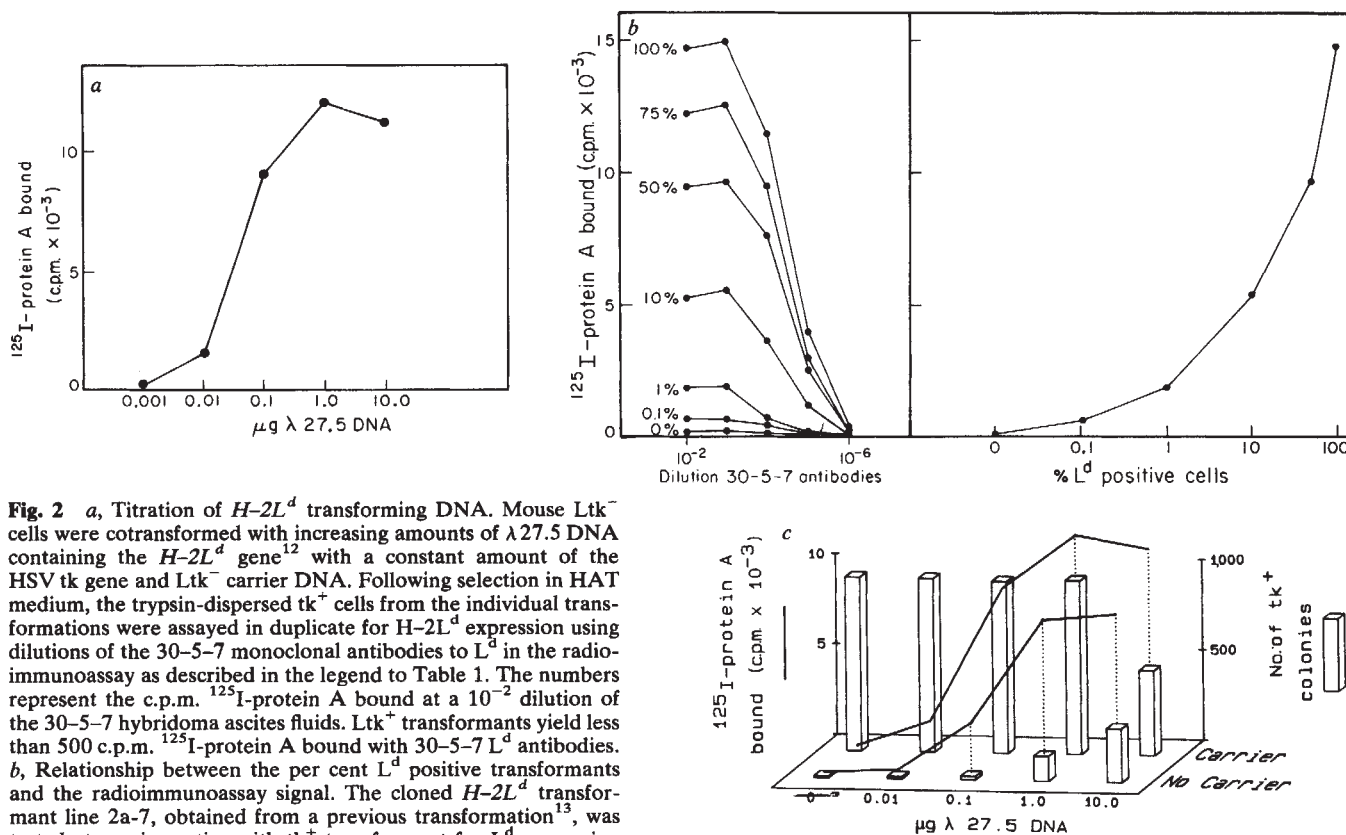


Fig. 2 a, Titration of $H-2L^d$ transforming DNA. Mouse Ltk^- cells were cotransformed with increasing amounts of $\lambda 27.5$ DNA containing the $H-2L^d$ gene¹² with a constant amount of the HSV tk gene and Ltk^- carrier DNA. Following selection in HAT medium, the trypsin-dispersed tk^+ cells from the individual transformations were assayed in duplicate for $H-2L^d$ expression using dilutions of the 30-5-7 monoclonal antibodies to L^d in the radioimmunoassay as described in the legend to Table 1. The numbers represent the c.p.m. ^{125}I -protein A bound at a 10^{-2} dilution of the 30-5-7 hybridoma ascites fluids. Ltk^+ transformants yield less than 500 c.p.m. ^{125}I -protein A bound with 30-5-7 L^d antibodies. b, Relationship between the per cent L^d positive transformants and the radioimmunoassay signal. The cloned $H-2L^d$ transformant line 2a-7, obtained from a previous transformation¹³, was tested at varying ratios with tk^+ transformant for L^d expression in the radioimmunoassay with dilutions of 30-5-7 monoclonal antibodies to L^d . Left, radioimmunoassay of K2a-7 cells combined with tk^+ transformants. Right, plot of the numbers of c.p.m. of ^{125}I -protein A bound for a 10^{-3} dilution of 30-5-7 antibodies for each of the reconstituted samples (from left-hand panel) versus the per cent L^d positive transformants in each test sample. c, Relationship between the concentration of cotransforming DNA, carrier DNA and positive cotransformation. Mouse Ltk^- cells were cotransformed with a constant amount of HSV tk gene and increasing amounts of $\lambda 27.5$ ($H-2L^d$) DNA in the presence or absence of carrier DNA. The numbers of tk^+ colonies were monitored closely and counted before dispersion in trypsin. Each of the pooled transformant lines was assayed for L^d expression as described in the legend to Table 1. The numbers represent the average c.p.m. of ^{125}I -protein A bound for a 10^{-2} dilution of the 30-5-7 hybridoma ascites fluids. The lines represent the c.p.m. for transformants obtained with and without carrier. The bars represent the numbers of colonies obtained with and without carrier.

apparently complement the expression of truncated class I molecules by a mechanism which leads to the production of full-length products resembling those of intact class I genes. To test the generality of these observations, several truncated forms of the well-characterized L^d gene were created and tested for their ability to generate intact class I molecules expressing L^d determinants after transfection into mouse L cells.

Subcloned fragments can generate complete L^d -like molecules

Plasmid constructs were prepared containing varying amounts of coding and flanking sequences from the L^d gene contained in clone λ 27.5^{12,14} (Fig. 4a). Plasmids 1500.1 and 1500.4 contain, in opposite orientations, 5' flanking DNA, exons 1, 2, 3, introns 1, 2 and 3, and 30 base pairs (bp) of exon 4. Because the first and second external domains, encoded by exons 2 and 3 (see Fig. 1), are highly variable, we believe they contain most of the polymorphic serological determinants recognized by appropriate monoclonal antibodies. Plasmids 1700.1 and 1700.8 contain, in opposite orientations, the 2.2-kb *Bam*HI 5' fragment of the 27.5 gene which presumably lacks the promoter sequences as it contains only 50 bp 5' to the first exon, exons 1 to 3, introns 1 and 2, and most of 3. The 1800.2 and 1800.3 plasmid constructs contain the 3' 2.4 *Bam*HI fragment of λ 27.5 with the 3' portion of third intron, exons 4 to 8, introns 4 to 7, and 1.6 kb of 3' flanking DNA. Plasmid p1102 contains 5' flanking sequences spliced to the second intron ending 5 kb to the 3' end of exon 8. These constructs are depicted in Fig. 4a.

Mouse Ltk⁻ cells were transformed with varying amounts of each of these recombinant plasmids and the resulting tk⁺ transformants assayed for the expression of cell-surface L^d antigen in the radioimmunoassay (Fig. 4b). In all cases the radioimmunoassay signals are substantially lower than those obtained by cotransformation with the complete $H-2L^d$ gene (compare with Fig. 2a). It is important to stress that mice of the $H-2^k$ haplotype do not express L molecules presumably because they lack the corresponding allele of this gene. In addition, mouse L cells do not express class I molecules reactive with the monoclonal anti- L^d reagent even after transformation with homologous L-cell carrier DNA and the HSV tk gene alone (see Table 1).

Constructs 1500.1 and 1500.4 (exons 1 to 3) yielded significant numbers of L^d positive transformants approaching 10% of the tk⁺ transformants at 10 μ g of DNA (Fig. 4b). This frequency, estimated from the magnitude of the radioimmunoassay signal of the pooled tk⁺ transformants, was confirmed by screening individual clones for the expression of L^d molecules. This clonal analysis revealed that 1 out of 10 of the tk⁺ clones express L^d molecules at levels comparable with the clones obtained by transformation with the complete gene (see legend to Fig. 4b). Therefore, transformation with the truncated L^d gene produces transformants expressing high levels of this cell-surface antigen at a low frequency rather than a high percentage of cells which express low levels of L^d molecules. We cannot exclude the possibility that additional transformation events leading to the production of non-antigenic molecules occurs. The orientation of the 1500-derived inserts had little or no effect on the production of L^d positive transformants. The omission of carrier DNA in the transformation procedure reduced the overall production of tk⁺ transformants but otherwise did not affect the frequency of L^d positive cotransformation (see transformation with p1500.4 DNA in Fig. 4b). This observation eliminates the possibility that the Ltk⁻ carrier DNA contributes the necessary coding information to the p1500 DNAs for productive transformation.

Transformation with DNA from p1102 (exons 3 to 8), which lacks the information for the promoter, leader peptide and first external domain, also generated L^d positive transformants at frequencies slightly lower but comparable with transformation with p1500.1 and p1500.4 DNAs (Fig. 4b). Therefore, both 5' and 3' coding information apparently can be complemented by transformation.

Transformation efficiencies were reduced to low but detectable levels using p1700.1 or p1700.8 DNAs (exons 1 to 3). We assume that in the case of these constructs, lacking both appreciable 5' flanking sequences (that is, a promoter) and exons 4 to 8, the 5' flanking sequences and exons 4 to 8 must be regenerated to form a complete $H-2L^d$ gene. Such a 'double event' may occur less frequently than the complementation of a single 5' or 3' region alone.

Transformation with p1800.2 or p1800.3 DNAs which contain exons 4 to 8 of the $H-2L^d$ gene, did not generate serologically detectable $H-2L^d$ positive transformants at observable frequencies. Thus, the data are consistent with the assumption

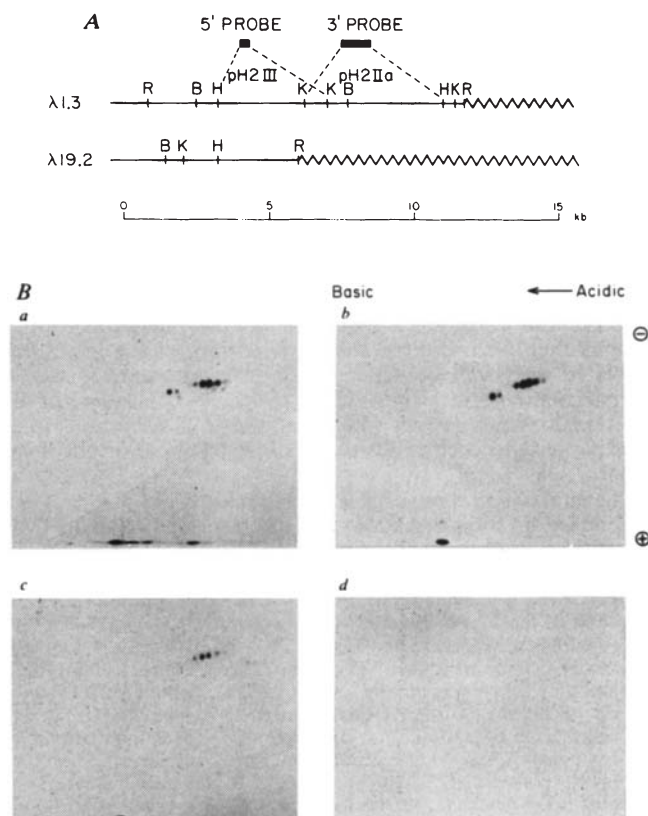


Fig. 3 A, restriction maps of λ recombinant DNA clones containing an $H-2K^d$ gene and the 19.2 gene. DNA from the λ clone 1.3, previously demonstrated to contain a functional $H-2K^d$ gene by transformation¹³, and the λ 19.2 gene were mapped by single and double digestion with the enzymes indicated below as previously described⁹. The radiolabelled probes from the nick-translated 5' probe pH-2III and the 3' probe pH-2IIa¹⁰ were used to detect fragments containing coding sequences homologous to the class I genes. The presence and orientation of specific coding regions on the restriction fragments were similarly determined using sequencing subclones to the λ 27.5 ($H-2L^d$) gene¹⁴. The fragments to which the class I cDNA probes hybridized are indicated by the dashed lines. The pH-2III probe hybridized to the *Hind*III-*Eco*RI fragment of λ 19.2 aligned below the hybridizing fragment of the λ 1.3 gene (dashed lines). No hybridization with either probe was observed with other regions of the 19.2 insert. The pH-2III probes contain the sequences encoding amino acids 63 to 160 of a transplantation antigen and pH-2IIa information for amino acids 167-352. The solid lines indicate the insert DNA, and the wavy lines the λ vector DNA. R = *Eco*RI, B = *Bam*HI, H = *Hind*III, and K = *Kpn*I. The 3' *Eco*RI site marks the position of the linker fragments used in the construction of the library. B, fluorographs of two-dimensional gels of $H-2K^d$ immunoprecipitates from BALB/cJ splenic lymphocytes (a), K7-65 transformants of the complete, normal $H-2K^d$ gene (b), cloned transformants (80-5a) of the λ 19.2 truncated gene (c), tk⁺ transformants (d) prepared for electrophoretic analysis as previously described¹². The major species migrated to the 45,000-MW region of the gel in the second dimension. β_2 -Microglobulin appears at the bottom of the fluorographs.

that transformation with DNA containing exon 3 (that is, p1500 or 1102 DNA) is sufficient for cotransformation of mouse cells to the L^d positive phenotype as defined with two separate monoclonal antibodies used in these assays (see legend to Fig. 4b). These observations suggest that these epitopes may be encoded by sequences, which in combination with the appropriate transformation elements, may be expressed. Alternatively, the observations suggest that the sequences required for the expression of the L^d serological determinant are probably contained on the second external domain.

Expression of complete L^d molecules by cells transformed with the truncated gene does not appear to be limited to the primary transformation event. L^d negative clones derived from transformation with 1500 DNA acquire at low frequency the L^d positive phenotype after several passages in culture while tk^+ control cells remain L^d negative (data not shown). This phenomenon is currently being investigated.

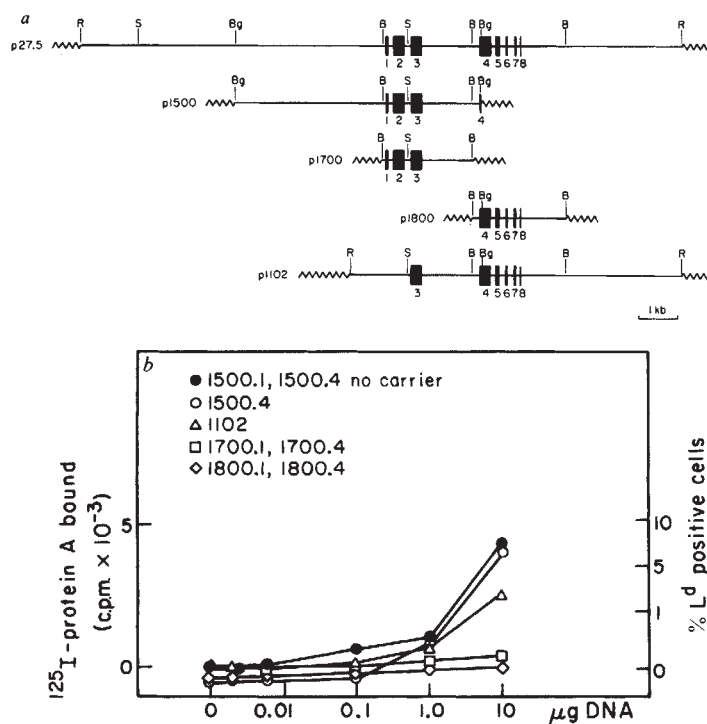
Truncated class I genes generate normal class I polypeptides

The L^d molecules produced by one cloned transformant (obtained with 1500.4 DNA without carrier) were compared with normal L^d products by electrophoretic analysis. Figure 5 shows that the two-dimensional gel pattern of the molecules isolated by immunoprecipitation from these cloned cells is superimposable with those of L^d molecules from spleen cells or transformants derived with the complete λ 27.5 gene despite the minor differences in the relative intensities of the individual spots.

Several observations suggest that the class I-like polypeptides produced by transformation with the truncated L^d and λ 19.2

genes are hybrid class I products. First, in both cases transformation yields products of 45,000 MW on denaturing gels, precisely the size of class I polypeptides (Figs 3b, 5). Second, the truncated class I genes λ 19.2 and 1500.4, lack a fourth exon encoding the third external domain and would therefore be unable to encode products associated with β_2 -microglobulin (Figs 3a, 4a). However, these truncated genes appear to have acquired a fourth exon since the gene products are expressed on the cell surface in association with β_2 -microglobulin (Figs 3b, 5). Third, the charge distributions of the class I-like polypeptides derived from transformation with the truncated genes are remarkably similar to the gene products expressed by lymphocytes (Figs 3b, 5). Two-dimensional gel electrophoresis can readily distinguish different class I genes and their allelic forms (compare the pattern and charge distributions of L^d and K^d molecules in Figs 3b and 5). A comparison of the sequences of the L^d and D^b molecules shows that their C-terminal regions (residues 167–350) exhibit virtually the same charge (see refs 14, 17, 18). This observation suggests that the charge on the C-terminal regions of class I molecules is highly conserved, and that the two-dimensional gel pattern of a class I molecule may be largely determined by the external domains as a result of glycosylation and sequence differences. Therefore, the fact that the charge heterogeneity is virtually indistinguishable for the L^d and p1500.4 gene products and for the K^d and λ 19.2 products suggests that the donor 3' regions of the hybrid genes must be derived from class I or highly related DNA sequences. Thus, the size, expression on the cell surface with β_2 -microglobulin, and the two-dimensional gel patterns all suggest that the class I-like polypeptides produced by transformation of mouse L cells appear to be hybrid class I molecules encoded by truncated class I genes which have acquired the domains encoded by deleted 3' exons from endogenous class I genes.

Fig. 4 a, Schematic representation of the exon-intron organization of the 27.5 L^d gene and its derivatives. Plasmid p27.5 was constructed by cloning the *EcoRI* DNA insert from λ 27.5 into the *EcoRI* site vector ACYC184³². p1500, 1700 and 1800 were made by subcloning *Bam*HI and *Bgl*II fragments of p27.5 into the phosphatase *Bam*HI-cleaved pBR322³³. Plasmid 1102 was constructed by partial *Sma*I digestion of p27.5. Preparations of competent *Escherichia coli* cells MC1061, transformation with hybrid plasmids, DNA preparations, and other related recombinant DNA techniques have been described previously³⁴. Plasmids were isolated from bacteria which were colony purified at least three times. Exon 1 encodes the signal or leader peptide, exons 2 to 4 the three external domains of class I molecules, exon 5 the transmembrane domain, and exons 6 to 8 encode the cytoplasmic domain as also depicted in Fig. 1. The wavy region represents the plasmid vector ACYC184 (clones 27.5 and 1102) or pBR322 (clones 1500, 1700 and 1800). Restriction enzyme sites used in the construction of the recombinant plasmids are shown by the abbreviations: S = *Sma*I, B = *Bam*HI, Bg = *Bgl*II, R = *Eco*RI. The coding information of the constructs was verified by restriction mapping and comparison with the published restriction map of the λ 27.5 gene¹⁴. b, Radioimmunoassay of mouse L cells transformed with the subcloned portions of the $H-2L^d$. Mouse Ltk⁺ cells were cotransformed with varying amounts of DNA from the $H-2L^d$ subclones, with or without carrier DNA as shown, and a constant amount of the HSV tk gene. Following selection in HAT medium, the tk^+ transformants were trypsin-dispersed and analysed for L^d expression in the radioimmunoassay with 30–5–7 antibodies. The numbers shown represent the average c.p.m. ¹²⁵I-protein A bound with a 10^{-2} dilution of the 30–5–7



hybridoma ascites fluid. Single lines are shown for the results obtained from transformations with DNAs contained in vectors in the opposite orientation since the numbers between such experiments varied by less than 5%. Transformations done in duplicate gave virtually identical results. Transformants derived from transfection with ACYC vector DNA alone yielded transformants with undetectable reactivity with the L^d antibodies (<300 c.p.m. ¹²⁵I-protein A bound). The L^d positive transformants also reacted with 28–14–8 L^d monoclonal antibodies³⁵ at levels comparable with those observed with 30–5–7 antibodies. However, the cotransformants yielded less than 500 c.p.m. ¹²⁵I-protein A bound with monoclonal antibodies specific for other $H-2^d$ class I antigens. The constructs are described in the text. The per cent L^d positive transformants indicated at the right was estimated from the data in Fig. 2b. Transformants produced by transfection with 10 μ g of p1500.4 DNA without carrier DNA were cloned as described in the legend to Table 1 and tested for L^d expression in the radioimmunoassay with 30–5–7 and 28–14–8 L^d antibodies. Of the 10 clones screened, only one exhibited significant numbers of c.p.m. bound with ¹²⁵I-protein A (that is, >17,000) following incubation with 30–5–7 or 28–14–8 antibodies, in agreement with the 7% (shown at right) estimated from the data in Fig. 3. This clone, designated K14–21.3, expresses normal levels of K^k molecules but reduced amounts of the D^k antigen.

Reduction in the expression of an L-cell class I gene

To determine whether transformation with the class I genes affects the expression of the endogenous L-cell class I genes, we monitored the levels of the K^k antigen expressed by various class I transformants. Cloned transformants derived from transfection with the $H-2K^d$, D^d and L^d genes were examined for the expression of host K^k molecules. Figure 6 shows that 10 clones of mouse Ltk^- cells or mouse L cells transformed only with the tk gene (Ltk^+) exhibit relatively constant amounts of K^k expression as measured by radioimmunoassay. In contrast, at least 9 out of 50 cloned transformants expressing K^d , L^d or D^d molecules express significantly lower levels of K^k molecules by radioimmunoassay with monoclonal antibodies. The levels of K^k molecules were reduced to the amounts detected on spleen cells of a $C3H(H-2^k \text{ haplotype}) \times DBA(H-2 \text{ haplotype}) F_1$ animal which expresses just a single K^k allele (see Fig. 6 legend). None of the clones tested lacked detectable K^k antigen but this might represent a less frequent event not observed given the number of clones screened. Preliminary analysis has revealed a similar pattern of reduction in the expression of D^k molecules (data not shown).

As there is a uniform distribution in the expression of $H-2^k$ molecules by the Ltk^- and Ltk^+ cells, the reduced $H-2^k$ expression does not result from clonal variation or transformation with tk and carrier DNA alone. In addition, the clones which express the lower levels of $H-2^k$ antigen do not express increased levels of the foreign gene products. In fact, transformation with the truncated $\lambda 19.2$ gene was accompanied by a reduction in K^k expression in mouse L cells despite the fact that almost all the clones screened were negative for the K^d phenotype (that is, only 1% of tk^+ 19.2 transformants express K^d molecules. It has been previously demonstrated that β_2 -microglobulin does not appear to be a limiting factor in the cell-surface expression of additional class I molecules by the transformants¹³. Therefore, these results are consistent with the inactivation of one of the two $H-2K^k$ genes during transformation by recombination with homologous class I

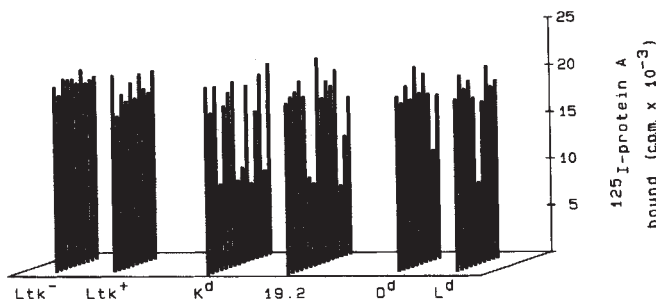


Fig. 6 $H-2K^k$ expression by the cloned class I transformants. The lines represent the average numbers of c.p.m. ($\times 10^{-3}$) of ^{125}I -protein A bound per 5×10^5 cells with a 10^{-1} dilution of 11.4 monoclonal hybridoma antibodies to K^k (ref. 36) molecules. The c.p.m. vary from the input c.p.m. by no more than 3.0% except for those clones considered to express abnormally low levels of K^k molecules. Standard deviations for the calculated means for normal values of low values varied less than ± 2.22 , respectively. Cells were transformed with 1.0 μ g of each of the DNAs using carrier DNA¹³ and subcloned in soft agar. Lymphocytes from F_1 animals of $C3H \times DBA$ origin demonstrated $\sim 50\%$ of the amount of ^{125}I -protein A bound compared with $C3H$ lymphocytes alone.

sequences, although we cannot exclude the possibility that post-transcriptional or translational regulation due to the insertion of class I sequences elsewhere contributes to this phenomenon. As Southern blot analysis of class I DNAs with cDNA probes cannot distinguish between the different class I genes in the mouse genome¹⁰, the donor sequences in the transformed cell cannot be localized to particular restriction fragments. Therefore, it is necessary to recloned the inserted elements to determine whether the loss of K^k expression is due to the insertion of donor sequences at that locus.

Discussion

DNA transformation can be divided into a complex series of molecular events involving transfection, integration and expression of exogenous DNA (for a review see ref. 19). It has been shown that integrated donor sequences are localized within concatamers formed between donor and carrier DNAs, which are randomly inserted in the genome of the recipient cell^{20,21}. A recent report suggests that recombination and even rearrangements between donor, donor and carrier, and possibly host DNAs occur, presumably during concatamer formation and integration^{22,23}. However, there is no direct evidence for recombination between homologous donor and endogenous host sequences such as that described for certain bacterial²⁴, lower eukaryotic²⁵ or viral systems^{26,27}. Moreover, it is not known whether the insertion of concatenates at specific sites in the chromosome of the recipient cell may be facilitated by homologous recombination. In bacteria²⁸ and yeast²⁹ the site of insertion of transforming DNA is dependent on the degree of homology between the exogenous and endogenous sequences. However, the size of the mammalian genome might reduce the effects of homology-dependent integration of transforming DNA by dilution of the target sequences.

Truncated class I genes containing 5' exons transform mouse L cells to yield complete hybrid class I molecules of the donor phenotype. These class I polypeptides appear to be recombinant in origin since they express serological determinants encoded by the donor DNA while at the same time exhibiting normal class I properties with regard to size, charge and the presence of a β_2 -microglobulin-binding domain supplied by the recipient cell. Thus, the evidence shows that transformation with truncated class I genes leads to the expression of intact hybrid class I genes. There are several possible explanations for this phenomenon. First, information from the donor and host

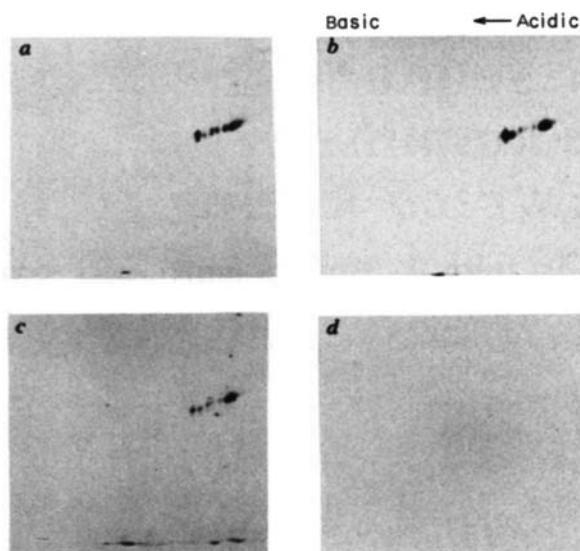


Fig. 5 Fluorographs of two-dimensional gels of $H-2L^d$ molecules. $H-2L^d$ immunoprecipitates from $\lambda 27.5$ transformants 2a-7 (a), BALB/cJ splenic lymphocytes (b), p1500.4 transformant clone 14-21.3 (c) (see Fig. 5 legend), and tk^+ transformants (d) were analysed on two-dimensional gels as previously described¹². The major species immunoprecipitated in each case migrated in the second dimension to the 45,000 MW region of the gel. β_2 -Microglobulin appears at the bottom of the fluorographs.

sequences could be joined either at the RNA or protein levels. There do not appear to be any previously described mechanisms for such processes (for example, RNA recombination or protein fusion). Second, these results also might involve the activation of endogenous class I genes in the host genome which encode products of the donor haplotype. This seems unlikely given that each strain of mouse appears to contain limited numbers of class I genes¹⁰ which presumably encode distinct transplantation antigens characteristic of each haplotype. Finally, the simplest and perhaps most plausible explanation is that homologous recombination or gene conversion occurs between the donor and host class I genes leading to the expression of hybrid class I polypeptides. To test this hypothesis, the inserted elements will be cloned and sequenced to identify the 3' L-cell class I exons involved in the presumed recombination event. The isolation of putative recombinant class I genes is complicated by the fact that multiple donor sequences are inserted into the host genome and these inserted sequences are quite homologous to the multiple endogenous class I in the host genome.

The critical question, apart from the mechanism for this phenotypic observation, is its generality, specifically whether

these events are unique to the class I genes of the major histocompatibility complex. The genes encoding the transplantation antigens constitute a highly polymorphic multigene family, suggesting that this polymorphism, at least in part, may arise from extensive recombination⁹. Hence there may be special genetic mechanisms or sequences in the class I gene family to promote extensive recombination. This phenomenon could be restricted to certain multigene families such as the antibodies and the class I genes. Finally, the apparent recombination between donor and host genes may occur with most eukaryotic genes, but at levels requiring more sensitive methods of detection. If hybrid genes are produced between homologous donor and host sequences, this would facilitate future attempts to transfer genes to specific sites on chromosomes for the correction of genetic defects.

We thank Ms Marie Krempin for technical assistance with the transformations and maintenance of the cell lines, Drs Ted Hansen and David Sachs for supplying the monoclonal antibodies, Drs Ellen Kraig, Mitch Kronenberg and Barbara Wold for their helpful discussion, and Ms Bernita Larsh for preparation of the manuscript. This work was supported by grants from the NIH, including AI 19297 (M.M.).

Received 16 November; accepted 15 December 1982.

1. Gorer, P. A. *J. Path. Bact.* **47**, 231–252 (1938).
2. Gorer, P. A., Lyman, S. & Snell, G. D. *Proc. R. Soc. B* **135**, 499–505 (1948).
3. Zinkernagel, R. M. & Doherty, P. C. *Adv. Immun.* **27**, 51–177 (1980).
4. Shearer, G. M. & Schmitt-Verhulst, A. M. *Adv. Immun.* **25**, 55–91 (1977).
5. Vitetta, E. S., Poulik, N. D., Klein, J. & Uhr, J. W. *J. exp. Med.* **144**, 179–192 (1976).
6. Krage, M. S., Orr, H. T. & Strominger, J. L. *Cell* **18**, 979–991 (1979).
7. Lopez de Castro, J. A. *et al. Biochemistry* **18**, 5704–5711 (1979).
8. Coligan, J. E., Kindt, T. J., Uehara, H., Martinko, J. & Nathenson, S. G. *Nature* **291**, 35–39 (1981).
9. Steinmetz, M., Winoto, A., Minard, K. & Hood, L. *Cell* **28**, 489–498 (1982).
10. Steinmetz, M. *et al. Cell* **24**, 125–134 (1981).
11. Steinmetz, M. S. *et al. Cell* **25**, 683–692 (1981).
12. Goodenow, R. S. *et al. Science* **215**, 677–679 (1982).
13. Goodenow, R. S. *et al. Nature* **300**, 231–237 (1982).
14. Moore, K. W., Sher, B. T., Sun, H. Y., Eakle, K. A. & Hood, L. *Science* **215**, 697–682 (1982).
15. Evans, G. A., Margulies, D. H., Camerini-Otero, R. E., Ozato, K. & Seidmann, J. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1994–1998 (1982).
16. Mellor, A. L. *et al. Nature* **298**, 529–533 (1982).
17. Martinko, J. M. *et al. Biochemistry* **19**, 6188–6193 (1980).
18. Uehara, H., Coligan, J. E. & Nathenson, S. G. *Biochemistry* **20**, 5940–5945 (1981).
19. Pellicer, A. *et al. Science* **209**, 1414–1422 (1980).
20. Peruchio, M., Hanahan, D. & Wigler, M. *Cell* **22**, 309–317 (1980).
21. Robins, D. M., Ripley, S., Henderson, A. S. & Axel, R. *Cell* **23**, 29–39 (1981).
22. Anderson, R. A., Krakauer, T. & Camerini-Otero, D. R. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2748–2752 (1982).
23. Roberts, J. M. & Axel, R. *Cell* **29**, 109–119 (1982).
24. Radding, C. M. A. *Rev. Biochem.* **47**, 847–880 (1978).
25. Stahl, F. W. *Genetic Recombination* (Freeman, San Francisco, 1979).
26. Winocour, E. & Keshet, I. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4861–4865 (1980).
27. Varmus, H. E., Quintrell, N. & Ortiz, S. *Cell* **25**, 23–36 (1981).
28. Radding, C. M. *Cell* **25**, 3–4 (1981).
29. Hinne, A. H., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929–1933 (1978).
30. Szybalska, E. H. & Szybalski, W. *Proc. natn. Acad. Sci. U.S.A.* **48**, 2026–2034 (1962).
31. Ozato, K., Hansen, T. H. & Sachs, D. H. *J. Immun.* **125**, 2473–2477 (1980).
32. Chang, A. C. Y. & Cohen, S. J. *J. Bact.* **134**, 1141–1156 (1978).
33. Bolivar, F. *et al. Gene* **2**, 95–133 (1977).
34. Stroynowski, I., van Cleemput, M. & Yanofsky, C. *Nature* **298**, 38–41 (1982).
35. Ozato, K. & Sachs, D. H. *J. Immun.* **126**, 317–321 (1981).
36. Oi, V. T., Jones, P. P., Goding, J. W., Herzenberg, L. A. & Herzenberg, L. A. *Curr. Topics Microbiol. Immun.* **81**, 115–129 (1978).

Extrachromosomal circular copies of an 'inter-Alu' unstable sequence in human DNA are amplified during *in vitro* and *in vivo* ageing

Robert J. Shmookler Reis, Charles K. Lumpkin Jr, John R. McGill, Karl T. Riabowol & Samuel Goldstein

Departments of Medicine and Biochemistry, University of Arkansas for Medical Sciences, and Geriatric Research Education and Clinical Center, Veterans Administration Medical Center, Little Rock, Arkansas 72205, USA

A DNA sequence situated in the human genome between Alu-repeat clusters ('inter-Alu' DNA) is progressively amplified in extrachromosomal DNA, including covalently closed DNA circles, during serial passage of diploid fibroblasts. A single-size class of inter-Alu circles is also amplified in lymphocytes from nine old donors and yet is not detected in cells from eight young donors.

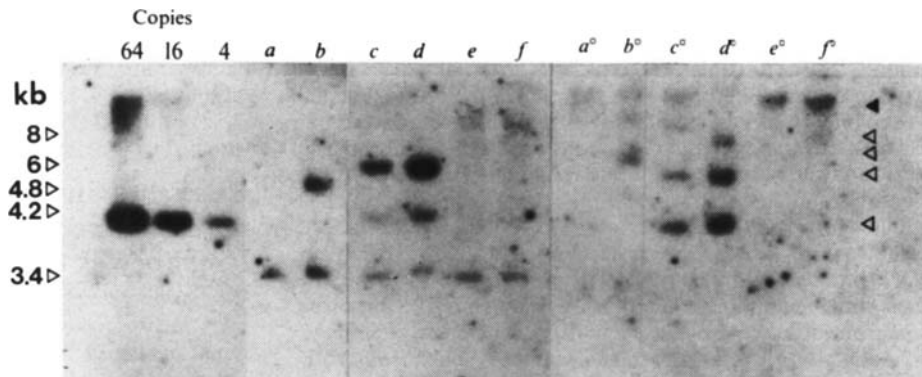
HIGHLY reiterated DNA sequences, including the 'EcoRI' family of centromeric tandem repeats¹ and three smaller repetitive classes (ref. 2, Fig. 3), become progressively depleted from the genome of diploid human fibroblasts during their replicative lifespan. The mechanism of loss is not known, but apparently involves small deletions of DNA¹, which could occur either by excision or through unequal recombination^{1,3}. Such DNA rearrangement events, although probably most common among

repetitious sequences³, cannot be individually identified in these families precisely because of their enormous repetitiveness (for example, the EcoRI repeat family comprises >300,000 copies per haploid genome). Study of the nature and extent of instability of highly repeated DNAs requires a unique or low-repetition sequence which is interspersed among them^{4,5}.

Calabretta *et al.*⁵ have recently isolated such a sequence from a 15-kilobase pair (kbp) cloned insert of human DNA which

Fig. 1 Integrated and extra-chromosomal *inter-Alu* copies in the DNA of three fibroblast strains, at early and late passage. DNA samples were from: strain A2, maximum lifespan 62 mean population doublings (MPD), at 17 MPD (a) and 51 MPD (b); strain DS, maximum lifespan 60 MPD, at 21 MPD (c) and 51 MPD (d); strain TM, maximum lifespan 52 MPD, at 15 MPD (e) and 49 MPD (f). Lanes a° – f° contain undigested DNA samples corresponding to a–f as above; that is, lane a° = lane a without enzyme digestion, and so on. Total high molecular weight DNA, visualized by ethidium bromide/UV fluorescence, either failed to enter the gel or ran at the minimum mobility for DNA ($\blacktriangle$). Hybridization to extrachromosomal DNA bands in advance of this is indicated by open triangles ($\triangle$).

Methods: DNA samples, cleaved with *Bam*HI (a–f) or undigested (a° – f°), were electrophoresed in 1.5% agarose gel (16 h at 2 V cm^{-1} , in 40 mM Tris-acetate, 4 mM NaOAc, 1 mM EDTA, pH 7.9) and transferred to a nitrocellulose filter (Schleicher & Schuell, BA85)³⁰. They were then hybridized to *inter-Alu* ³²P-DNA probe (10^6 c.p.m. ml^{-1} , $2\text{--}6 \times 10^8$ c.p.m. per μg , labelled by nick-translation³¹) and autoradiographed with Kodak XAR film at -70°C . Approximate fragment sizes in kilobase pairs (kbp), as indicated, were determined from PM2/*Hind*III fragments. Loads were 2 μg per 0.5-cm slot. Hybridization standards, corresponding to 4, 16 and 64 copies of *inter-Alu* per cell, were prepared by serial dilution of *Eco*RI-digested plasmid pAH15A.



includes at least 10 copies of the *Alu* repeat. They excised and subcloned a 0.8-kbp region situated between clusters of four and three *Alu* repeats. Hybridization of this '*inter-Alu*' probe against DNA from human tissues indicated moderate repetitiveness, up to 50 copies per cell. Considerable polymorphism was observed between tissues of a single donor, and between lymphocytes of different leukaemic and normal donors, in both the *inter-Alu* banding pattern and copy number⁵. The restriction fragment length polymorphism and most of the copy number variation seemed to be due to circular extrachromosomal molecules containing this sequence. A 4.8-kbp band, which constituted the principal-size class of circles hybridizing to *inter-Alu* probe in many tissues, also contained *Alu* repeats but differed from genomic copies of *inter-Alu* with respect to mapping of adjoining restriction sites⁵. Because of the apparent instability of the *inter-Alu/Alu* repeat cluster, we have monitored it as an index of genomic plasticity during the limited replicative lifespan of diploid human fibroblasts ('*in vitro* ageing'), and during the lifespan of individual cell donors ('*in vivo* ageing').

Amplification of discrete extrachromosomal species during *in vitro* ageing

We examined six normal diploid fibroblast strains (three of these in duplicate expansions) during the course of their replicative lifespans. DNA samples, both undigested and following cleavage with *Bam*HI restriction endonuclease, were electrophoresed in agarose gels, Southern-blotted on nitrocellulose filters⁶ and hybridized to *inter-Alu* probe. Results for three fibroblast strains at early and late passage are shown in Fig. 1. In each case after *Bam*HI digestion the genomic (integrated chromosomal) *inter-Alu* band appeared at 3.4 kbp (Fig. 1a–f), as expected from the restriction map of the original 15-kbp genomic DNA clone⁵. The genomic copy numbers varied between cell strains, ranging from two to six per cell, but generally remained essentially constant ($\pm 40\%$) for any given strain during serial passage.

Four of the six cell strains showed discrete additional DNA bands appearing or increasing at late passage (Fig. 1a–f). These are believed to represent DNA species which are extrachromosomal (that is, not integrated into large chromosomal DNA; see below and ref. 5) and circular, although circularity was not proven in each instance. They varied widely in size, bands of 1.6, 4.2, 4.8, 6 and 8 kbp being observed following *Bam*HI digestion. The extent of amplification also varied between cell strains, increasing from 0 to 4 copies per cell in strain A2 (Fig. 1a, b), from 6 to ~ 20 copies per cell in strain DS (Fig. 1c, d), and from 1 to 4 copies per cell in strain TM (Fig. 1e, f).

Restriction cleavage was complete in each case, as further increase in enzyme:DNA ratio or incubation time had no effect on the band pattern (data not shown).

Note that all fibroblast strains reported here were derived by us from forearm biopsies of healthy donors^{7,8} and retain the diploid karyotype even at late passage (ref. 1 and unpublished data). They have been maintained continuously in our laboratory, free of mycoplasma (unambiguously negative by both fluorescent staining and uridine/uracil incorporation assays^{9,10}) without any use of antibiotics, which may lead to amplification of circular extrachromosomal DNA^{11,12}. Control hybridizations against pure mitochondrial DNA showed no homology to *inter-Alu* probe (data not shown), indicating that the circular DNA molecules in question were not derived from the mitochondrial genome. The chromosomal/genomic bands served as internal controls for hybridization efficiency, as they generally remained at constant intensity (or altered slightly) whereas the extrachromosomal bands increased at late passage (Fig. 1a–f).

Amplified *inter-Alu* sequences are in extrachromosomal covalently closed circles

The extrachromosomal nature of these variable, passage-dependent bands was confirmed by hybridization of *inter-Alu* probe to undigested total cell DNA (Fig. 1a $^{\circ}$ –f $^{\circ}$). Whereas the bulk of the DNA ran as a diffuse region near the top of the gel ($\blacktriangle$ in Fig. 1a $^{\circ}$ –f $^{\circ}$), several discrete hybridizing bands migrated faster than this (open triangles in Fig. 1a $^{\circ}$ –f $^{\circ}$). In addition, hybridization to DNA in the high molecular weight region ($\blacktriangle$) indicated integrated chromosomal copies of *inter-Alu*. Reconstruction experiments have shown that added circular DNA of probe size (plasmid pBR322, 4.4 kbp) does not become trapped by high molecular weight DNA, while ³²P-labelled DNA probes for endogenous genes such as β - and γ -globin sequences² hybridized exclusively to this chromosomal DNA region in undigested samples (data not shown). Thus, the observed difference in mobility between 'chromosomal' and 'extrachromosomal' sequences would not have arisen artefactually.

For each new *inter-Alu* band in *Bam*HI-digested DNA, there were generally two bands in undigested DNA—presumably corresponding to nicked and closed circular forms⁵. Both the number and intensity of discrete bands in undigested DNA were well correlated with non-genomic bands in restriction-digested DNA (compare left and right sides of Fig. 1). Thus, the increase in *Bam*HI bands (other than 3.4 kbp) at late passage was reflected in an increased number and intensity of discrete hybridizing bands in undigested DNA, running ahead of the high molecular weight linear DNA region (Fig. 1a $^{\circ}$ –f $^{\circ}$).

The close agreement between these two assays, using cleaved and uncleaved DNA, indicates that most, if not all, of the amplified *inter-Alu* copies are extrachromosomal.

To demonstrate the presence of amplified *inter-Alu* sequences in covalently closed circular DNA (cccDNA), chromosomal nucleoprotein was precipitated¹³ from nuclei of A2 fibroblasts at early and late passage, and the supernatants were centrifuged to equilibrium in CsCl gradients containing 200 $\mu\text{g ml}^{-1}$ ethidium bromide¹⁴. Hybridization to fractions from these gradients (Fig. 2a, b) indicated *inter-Alu* hybridization to cccDNA molecules (density 1.59 g ml^{-1}) as well as to nicked circles and linear molecules (density 1.55 g ml^{-1}). Moreover, extrachromosomal *inter-Alu* sequences were greatly enriched (>50-fold, per μg of DNA) in the cccDNA peaks from both

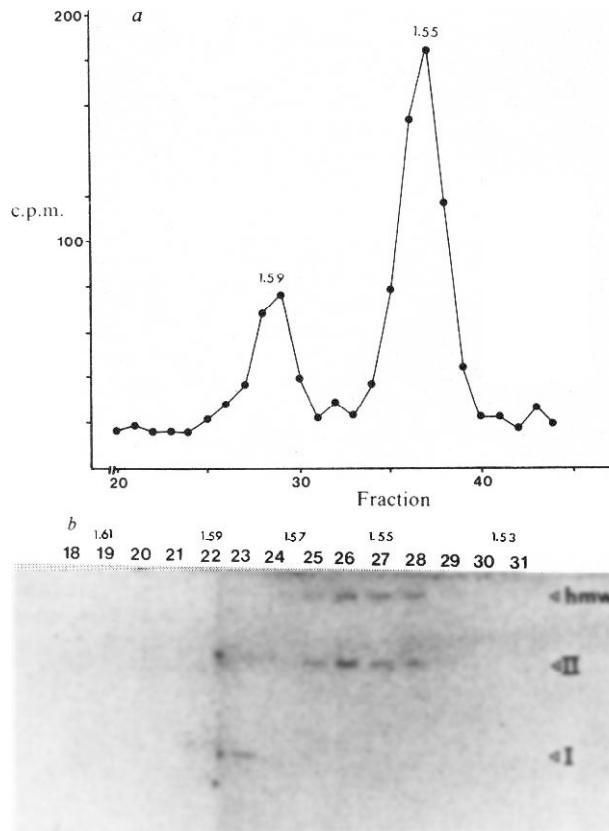


Fig. 2 Identification of *inter-Alu* sequences in cccDNA. *a*, ^3H -DNA banded in CsCl/ethidium bromide, isopycnic gradient. *b*, Hybridization of *inter-Alu* ^{32}P -DNA probe to fractions from a gradient (similar to that in *a*) containing A2 DNA at 60 MPD. **Methods:** In *a*, very late-passage A2 fibroblasts (60 MPD) were labelled for 6 days (~ 1 MPD) with 1 $\mu\text{Ci ml}^{-1}$ ^3H -thymidine. Cells were collected, rinsed twice and lysed in hypotonic medium using a Dounce homogenizer. Nuclei were rapidly pelleted through 0.5 M sucrose and lysed in alkaline SDS; chromosomal nucleoprotein was precipitated after neutralization¹³ and the supernatant was concentrated by ethanol precipitation, for isopycnic centrifugation CsCl (density 1.57 g ml^{-1}) containing 200 $\mu\text{g ml}^{-1}$ ethidium bromide. Fractions at densities 1.59 and 1.55 g ml^{-1} , as indicated, correspond to cccDNA and linear/nicked-circular DNA, respectively³². ^3H -DNA c.p.m. per 10- μl aliquot are shown. In *b*, following removal of ethidium bromide by isopropanol extraction, 2- μl aliquots from 0.1-ml fractions were diluted to 10 μl in gel buffer and electrophoresed on a 1.5% agarose gel. DNA was blotted onto a nitrocellulose filter⁶, which was hybridized to *inter-Alu* probe as described in Fig. 1 legend. Densities of fractions, determined from refractive indices, are indicated above the fraction numbers. Positions are marked for *inter-Alu* form I (cccDNA) and form II (nicked circular DNA) bands; 'hmv' indicates high molecular weight DNA hybridization which may represent *inter-Alu* sequences in large linear molecules or in large nicked circles (concatenates or oligomers).

early- and late-passage cells, but were 8–10-fold more abundant in cccDNA of cells at late passage (Fig. 3b, d, f, h) than at early passage (Fig. 3a, c, e, g). These cccDNA molecules ran predominantly at 4.8 kbp after cleavage with *Hind*III (*a*, *b*), *Eco*RI (*c*, *d*) or *Bam*HI (*g*, *h*), consistent with a single cleavage site (or cluster) for each enzyme, on a circular molecule.

Extrachromosomal copies of the *inter-Alu* sequence increase dramatically in lymphocytes during *in vivo* ageing

If *in vitro* ageing of cells is accompanied by a variable increase in extrachromosomal copy number for the *inter-Alu* sequence, might ageing of the organism have a similar effect? To study the effects of *in vivo* ageing, we prepared DNA from peripheral lymphocytes of eight 'young' normal donors, age 21–31 yr, and nine 'old' normal donors, age 61–91 yr. All donors were in apparent good health and were living at home at the time of blood collection; their nutritional status is being investigated. Lymphocytes were prepared from fresh blood samples by standard procedures¹⁵ and their DNA was purified¹. Samples of total lymphocyte DNA, either undigested or digested with *Bam*HI endonuclease, were electrophoresed, Southern-blotted and probed for *inter-Alu* sequences. Results for four young and four old donors are shown in Fig. 4. All samples contained the genomic (3.4 kbp) *inter-Alu* band, for which the copy numbers (estimated by comparison with hybridization standards, serial dilutions of *Eco*RI-cleaved *inter-Alu* plasmid DNA) in both young and old individuals ranged from 2 to 30–40 copies per cell. Several-fold variation in the chromosomal *inter-Alu* copies had been observed by Calabretta *et al.* (ref. 5 and personal communication) within a smaller sample of normal leukocyte donors, but the extent of population heterogeneity we have observed was unexpected.

No *inter-Alu* bands other than the genomic (3.4 kbp after *Bam*HI) were seen in DNA samples derived from any of the eight young donors' lymphocytes (Fig. 3a–d). In striking contrast, all nine old donors' samples had a single additional band at 4.8 kbp (Fig. 4e–h), at levels of 4–40 copies per cell. This age-dependent *inter-Alu* band was evidently extrachromosomal because undigested DNA samples from old donors' (but not young donors') lymphocytes also showed two discrete bands of hybridization (<), in advance of the unresolved chromosomal

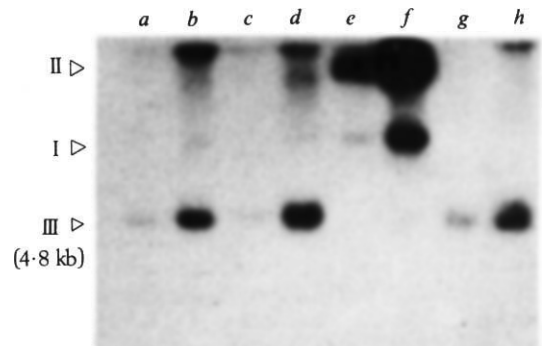
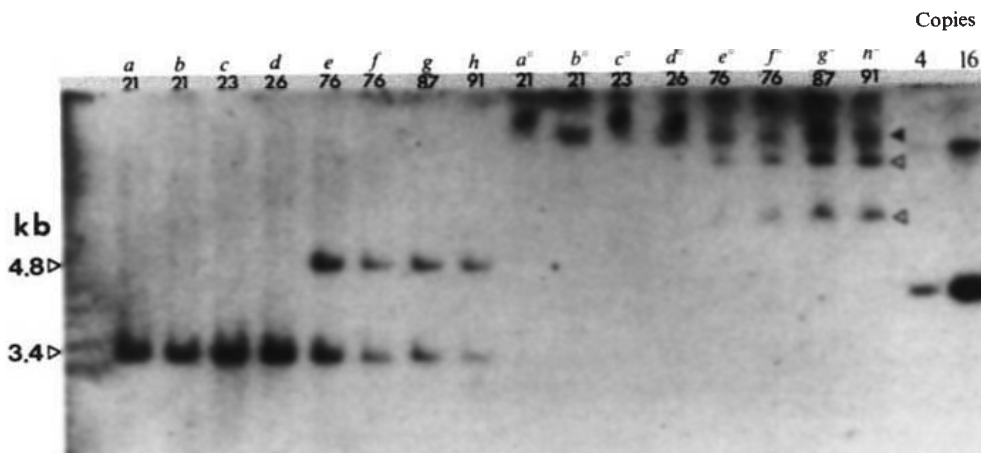


Fig. 3 Analysis of cccDNA from human fibroblasts. *a*, *b*, Digested with *Hind*III; *c*, *d*, digested with *Eco*RI; *e*, *f*, undigested; *g*, *h*, digested with *Bam*HI. Forms I (cccDNA), II (nicked circles) and III (linears) are indicated for the *inter-Alu* 4.8-kbp extrachromosomal molecules. A2 cccDNA at 28 MPD: lanes *a*, *c*, *e*, *g*; A2 cccDNA at 53 MPD: lanes *b*, *d*, *f*, *h*.

Methods: Extrachromosomal DNA from A2 fibroblasts at 28 and 53 MPD ($\sim 10^8$ cells each) was prepared by alkaline-SDS lysis of nuclei and centrifuged in CsCl/ethidium bromide as described in Fig. 2 legend. The gradient fractions corresponding to cccDNA were pooled, ethidium bromide was removed by isopropanol extraction and DNA was recovered by ethanol precipitation. These cccDNA preparations were electrophoresed in adjacent lanes of an agarose gel, Southern-transferred and probed with *inter-Alu*. DNA samples (0.1 μg) each represent 5×10^6 cells. Restriction bands of genomic length were not seen in these fractions.

Fig. 4 Integrated and extrachromosomal *inter-Alu* copies in human lymphocyte DNA from young and old donors. DNA samples (2 µg each) were from donors of the indicated ages (yr): a, 21; b, 21; c, 23; d, 26; e, 76; f, 76; g, 87; h, 91. Hybridization standards, on the right, correspond to 4 and 16 *inter-Alu* copies per cell. Lanes a° – h° correspond to lanes a–h; that is, lane a° = lane a without digestion, and so on. High molecular weight chromosomal DNA ran at or above the closed triangles (▲) and could be visualized by ethidium bromide staining and UV fluorescence, whereas extrachromosomal bands running in advance of this (◀) could be visualized only by hybridization to ^{32}P -labelled probe and autoradiography (e° – h°).

Methods: Lymphocyte DNA samples, digested to completion with *Bam*HI (a–h) or undigested (a° – h°), were electrophoresed, transferred to nitrocellulose and hybridized to *inter-Alu* ^{32}P -DNA probe (see Fig. 1 legend).



linear DNA (◀) (Fig. 4 e° – h°). These two bands, presumably corresponding to covalently closed circular DNA (cccDNA, faster band) and nicked circular DNA (slower band), were consistently found at the same mobilities in lymphocyte DNA samples from old donors. The close association between donor age and extrachromosomal *inter-Alu* copies, seen in *Bam*HI-digested DNA, was thus confirmed by this assay on undigested DNA.

Lymphocytes from two of the older donors were fractionated into T and B cells, and their DNAs were then isolated and probed for *inter-Alu* sequences. In both fractionations, the amplified *inter-Alu* circles were found predominantly in B lymphocyte DNA; indeed, the residual level of circles apparent in our T-cell preparations may be accounted for by B-cell contamination of this fraction (manuscript in preparation).

Discussion

Human¹¹ and monkey¹⁶ permanent cell lines have been reported to contain small polydisperse DNA circles, at least some of which include *Alu* interspersed repeat sequences^{5,16}. Calabretta *et al.*⁵ have found both chromosomal and extrachromosomal (putatively circular) forms of *inter-Alu* DNA in normal and cancerous human tissues, showing considerable inter-tissue and inter-individual polymorphism⁵. Our results corroborate those of Calabretta *et al.*⁵, in particular with respect to the extrachromosomal and closed-circular status of *inter-Alu* bands other than the predicted genomic band. In addition to extensive population heterogeneity of genomic *inter-Alu* copy numbers, we have found a remarkable association between *inter-Alu* circular forms and senescence, both in terms of cellular replicative lifespan (*in vitro* ageing) and organism lifespan (*in vivo* ageing). These two model systems for the study of senescence have both differences and similarities which deserve emphasis. Neurohumoral changes *in vivo* (including hormonal and nutritional variables) may be involved in the age correlation for lymphocytes, whereas *in vitro* cell cultures have the distinct advantage of isolation from such variables. In each case the amplification of extrachromosomal *inter-Alu* copies accompanies senescence, suggesting a common molecular concomitant of ageing in both systems. Despite the attractiveness of this parallel, we would caution that the two systems might 'senesce' for different reasons, and could conceivably amplify *inter-Alu* circles by different mechanisms. Thus far, the principle difference we have noted is that amplified *inter-Alu* molecules are polymorphic with respect to size in fibroblasts, but not in lymphocytes.

As both lymphocytes and the fibroblast cultures are heterogeneous mixtures of cells, it is possible that each contains minor subpopulations of circle-bearing cells which come to dominate the population during senescence. In the case of

lymphocytes, our evidence indicates that *inter-Alu* circles arise primarily in B rather than T cells. Because the proportion of B lymphocytes in human peripheral blood does not alter with age^{17,18}, the age-dependent appearance of *inter-Alu* 4.8-kbp circles cannot be attributed to generalized B lymphocyte hyperplasia. Fibroblasts, however, can undergo 'clonal succession' during *in vitro* culture^{8,19,20}. We have found initial evidence of interclonal heterogeneity for *inter-Alu* circle number, but none of the five clones assayed at early passage had as many *inter-Alu* circles as the mass cultures at the same doubling level. If the increase in *inter-Alu* circles were due to shifts in the clonal composition of the fibroblast mass culture, some clones at early passage must contain at least as many circles per cell as are contained in the late-passage mass culture. Failure to find such clones would argue against clonal succession as a sole explanation for *inter-Alu* amplification. Further investigations are in progress to pursue this question, and to ascertain (1) whether circle numbers increase within each clone, (2) whether specific clones rich in circles plate more efficiently or proliferate more rapidly than others, and (3) whether an increasing proportion of clones acquire high numbers of circles at late passage.

Mechanisms for generating *inter-Alu* circles

Circular molecules could in principle arise by excision (via recombination) from linear chromosomal sequences, or by autonomous replication of preexisting circles. Autonomous replication has not been directly demonstrated for these extrachromosomal circular molecules, but three lines of evidence support its occurrence. First, human diploid fibroblasts at very late passage, undergoing little chromosomal DNA synthesis, can nevertheless incorporate ^3H -thymidine into cccDNA (identified by buoyant density, 1.59 g ml^{-1} , in CsCl /ethidium bromide gradients), as shown in Fig. 2a. Second, both fibroblasts and normal lymphocytes can increase *inter-Alu* circle numbers to levels greatly exceeding their genomic copy numbers (Fig. 1 and refs 5, 21); this clearly cannot occur by excision alone, but would require autonomous replication of circles or perhaps a 'chromosome-copy' mechanism^{22,23}. As circular *inter-Alu* forms have restriction site maps which differ from the genomic map⁵, it is difficult to envisage their origin from an integrated chromosomal copy, without also invoking rearrangement or mutation of circles and autonomous replication. Third, the selective amplification of circular species with variable sizes in different fibroblast strains (Fig. 1) and in different tissues or individuals⁵ would be most easily explained by autonomous replication of a specific circular form in each instance.

Replicate expansions of three fibroblast strains, derived from separate aliquots of cells in liquid N_2 storage, produced very similar passage-dependent increases in *inter-Alu* circles. For

two strains, the amplified size class remained the same; in the third strain, however, one expansion produced a single prominent extrachromosomal *inter-Alu* band while the second expansion produced three size classes. This suggests that the different sizes of extrachromosomal circles appearing in different cell strains may be due to stochastic events in culture rather than donor genotype alone.

Another hypothesis would be that circular species containing *inter-Alu* sequences arise via homologous recombination²⁴ or gene conversion²³ between genomic *inter-Alu* copies and endogenous circular species, and are subsequently amplified during ageing *in vitro* or *in vivo*. A further possibility, that such circular forms are infectious and transmitted to cells or individuals in a time-dependent fashion, may be unlikely in view of the different circular sizes appearing in several late-passage fibroblast cultures grown concurrently. That is, the chance of infection in all four cultures, each with different viral species, seems remote. In this context it is also noteworthy that none of the aged lymphocyte donors had previous contact with members of our laboratory.

Analogy of the *inter-Alu/Alu* cluster to prokaryote and eukaryote transposable elements

Alu repeat units in human DNA possess many properties characteristic of prokaryotic and eukaryotic insertion sequences²⁵⁻²⁸: most units are flanked by short direct repeats; they contain promoter sequences for an RNA polymerase (pol III);

they are found both in circular DNA molecules^{5,16} and scattered within the genome; and they have evidently inserted into other genomic sequences, forming flanking direct repeats by duplication of a segment of the pre-insertion site sequence²⁹.

The *inter-Alu* sequence together with adjacent *Alu* repeat units is structurally analogous to several prokaryote transposons, which consist of two identical insertion sequences flanking a central region that generally encodes known gene products²⁵. Several transposable elements which have been characterized in eukaryotes share this 'transposon-like' structure: for example, *Ty1* in yeast, and *copia*, *412* and *297* in *Drosophila*^{25,26}. Although the case for transposability of *Alu* itself is reasonably strong²⁷⁻²⁹, it must be emphasized that transposition of the *inter-Alu* sequence remains to be demonstrated.

Thus, although many aspects of the polymorphic *inter-Alu* sequence have yet to be resolved, their striking association with both *in vitro* and *in vivo* ageing poses an intriguing problem. Regardless of whether they are implicated as causative agents in senescent loss of adaptive function, or instead prove to be a secondary phenomenon perhaps reflecting the more general loss of genomic stability during the lifespan, they reveal a remarkable potential for structural flux in eukaryotic genomes.

We thank Dr Bruno Calabretta and Dr Grady Saunders for the *inter-Alu* plasmid (pH15A) and many valuable discussions, and Elena J. Moerman, Richard Jones, Monica Gonzalez and Richard Denton for technical assistance. This research was supported by grants from the NIH (AG-03314, AG-03787) and the Veterans Administration.

Received 22 April; accepted 15 November 1982.

- Shmookler Reis, R. J. & Goldstein, S. *Cell* **21**, 739-749 (1980).
- Shmookler Reis, R. J. & Goldstein, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3949-3953 (1982).
- Harley, C. B., Shmookler Reis, R. J. & Goldstein, S. *J. theor. Biol.* **94**, 1-12 (1982).
- Szostack, J. W. & Wu, R. *Nature* **284**, 426-430 (1980).
- Calabretta, B., Roberson, D. L., Barrera-Saldana, H. A., Lambrou, T. P. & Saunders, G. F. *Nature* **296**, 219-225 (1982).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Singal, D. P. & Goldstein, S. *J. clin. Invest.* **52**, 2259-2263 (1973).
- Harley, C. B. & Goldstein, S. *J. cell. Physiol.* **97**, 509-516 (1978).
- Schneider, E. L. *Meth. Cell Biol.* **10**, 261-276 (1975).
- Schneider, E. L. & Stanbridge, E. J. *Meth. Cell Biol.* **10**, 277-290 (1975).
- Smith, C. A. & Vinograd, J. *J. molec. Biol.* **69**, 163-178 (1972).
- Stanfield, S. & Helinski, D. R. *Cell* **9**, 333-345 (1976).
- Birnboim, H. C. & Doly, J. *Nucleic Acids Res.* **7**, 1513-1523 (1979).
- Smith, C. A., Jordan, J. M. & Vinograd, J. *J. molec. Biol.* **59**, 255-272 (1971).
- Boyum, A. *Scand. J. Immun.* **5**, Suppl 5, 9-15 (1976).

- Krolewski, J. J., Bertelsen, A. H., Humayun, H. Z. & Rush, M. G. *J. molec. Biol.* **154**, 399-415 (1982).
- Gupta, S. & Good, R. A. *J. Immun.* **122**, 1214-1219 (1979).
- Gillis, S., Kozak, R., Durante, M. & Weksler, M. *J. clin. Invest.* **67**, 937-942 (1981).
- Salk, D., Au, K., Hoehn, H., Stenchever, M. R. & Martin, G. M. *Cytogenet. cell. Genet.* **30**, 108-117 (1981).
- Wojtyk, R. I. & Goldstein, S. *J. Cell Biol.* **95**, 704-710 (1982).
- Shmookler Reis, R. J., Lumpkin, C. K. Jr, McGill, J. R., Riabowol, K. T. & Goldstein, S. *Cold Spring Harb. Symp. quant. Biol.* (in the press).
- Brown, D. D. & Blackler, A. W. *J. molec. Biol.* **63**, 75-83 (1972).
- Baltimore, D. *Cell* **24**, 592-594 (1981).
- Scherer, S. & Davis, R. W. *Science* **209**, 1380-1384 (1980).
- Calos, M. P. & Miller, J. H. *Cell* **20**, 579-595 (1980).
- Green, M. M. A. *Rev. Genet.* **14**, 109-120 (1980).
- Jagadeeswaran, P., Forget, B. G. & Weissman, S. M. *Cell* **26**, 141-142 (1981).
- Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065-1070 (1982).
- Grimaldi, G. & Singer, M. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1497-1500 (1982).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184-1188 (1975).
- Bauer, W. & Vinograd, J. *J. molec. Biol.* **33**, 141-171 (1968).

LETTERS TO NATURE

Inflation and time asymmetry in the Universe

P. C. W. Davies

Department of Theoretical Physics, The University, Newcastle upon Tyne NE1 7RU, UK

The recently proposed inflationary Universe scenario¹⁻⁴ explains several of the mysteries of modern cosmology. I argue here that it also provides a natural explanation for the origin of time asymmetry ('time's arrow') in the Universe. The new feature which inflation injects into this long-standing problem is the temporary dominance of the cosmological term in the gravitational field equations, which acts as a sort of repulsive gravity. This term generates huge quantities of energy and radiation (or matter) entropy, while drastically reducing the entropy density of the gravitational field. It thus establishes a large gap between the radiation entropy and the gravitational entropy, which gravity is now trying to close.

Most discussions of the origin of 'time's arrow' appeal ultimately to cosmology⁵. There is a paradox that, according to the standard hot big-bang model, the material contents of the Universe began in a condition of thermodynamic equilibrium (primeval chaos), whereas the Universe today is highly ordered, and far from equilibrium. The apparent reduction in entropy implied seems to conflict with the requirement of the second law of thermodynamics, which demands that entropy should never decrease. The natural tendency is for the Universe to 'run down' towards the heat death, and the problem is to explain how it was 'wound up' in the first place.

The paradox is partially resolved once it is appreciated that the Universe is an open system, whereas the second law applies to closed systems. The openness involved is not, however, of a spatial nature. The crucial observation is that the local cosmological material is open to the external cosmological gravitational field, through which, by the general theory of relativity, one describes the expansion of the Universe.

Specifically, in the high-density primaeval phase (in a simplified, conventional model) the expansion, although rapid, was much slower than the relaxation rate for typical particle processes in the cosmological material. The structure of the

material therefore easily adjusted itself to the evolving circumstances, such as declining temperature, without lagging behind equilibrium conditions. Entropy was at a maximum consistent with the cosmological constraints.

After a few seconds, however, the situation changed dramatically as nucleosynthesis became entropically favoured, because the relaxation time for nuclear burning was very much longer than the expansion time scale of that epoch. The material thus got out of step with the expansion, and an entropy gap opened up. There was simply insufficient time for equilibrium to be established. Most of the material remained 'frozen' in the form of low-entropy hydrogen, which is now synthesizing heavy elements in stars in an attempt to restore equilibrium. Our present existence is due to the lag induced by the rapid cosmological expansion in the first few minutes which had the effect of 'winding up' the Universe. Rather than the entropy decreasing, what has happened is that entropy has actually increased, but the maximum possible entropy also increased—and faster—due to the changing cosmological constraints implied by the expansion of the Universe. This gap-opening mechanism has been made explicit in a recent model calculation⁶.

Invoking the expansion of the Universe provokes us to contemplate the entropic qualities of the cosmological gravitational field in particular, and gravity in general. The idea that there should be a 'gravitational entropy' has been much discussed^{5,7-9}, while Bekenstein¹⁰ and Hawking¹¹ supplied an expression for the entropy of a black hole.

Although we still lack a formulation for the entropy of an arbitrary gravitational field, some qualitative features of gravitational entropy are easily deduced. The tendency of self-gravitating systems to grow clumpy and inhomogeneous from smooth initial states is obviously irreversible, the black hole representing the equilibrium end state of this spacetime crumpling. This time-asymmetric tendency ought to receive expression within a generalized second law of thermodynamics⁷.

Accepting that a smooth gravitational field represents a low-entropy state, one arrives at a curious reversal of the foregoing conclusion pertaining to the cosmological material. Whereas the material could have started out in thermodynamic equilibrium, and thus did not need to be created in any very special state to account for the currently observed order in the Universe, the gravitational field apparently started out in an exceedingly improbable state. The exceptional uniformity of the Universe on the large scale testifies to the relatively smooth nature of the gravitational field in the *primaeval* phase.

This paradox has been emphasized by Penrose⁸, who poses the problem in the following way. If, instead of being distributed more or less evenly across the cosmos, the cosmological material were accumulated into a black hole, the entropy of the hole (using Hawking's formula¹¹) would be some 10^{30} times greater than the actually observed entropy of the Universe. Assuming the usual relationship between entropy and statistics, this gives odds of $\sim 10^{10^{30}}$ to one against the present gravitational arrangement happening by chance. Why, asks Penrose, when there are so many ways for the big bang to cough out black holes, did it actually disgorge smoothly distributed matter? How did the Universe manage to go bang in such an improbable way?

The argument applies not only to black holes, but also to the general isotropy and homogeneity of the Universe. With so many degrees of freedom available for expansion, why was all the explosive energy concentrated in a single dilatory mode of uniform distension? A highly turbulent, irregular Universe would have been much closer to thermodynamic equilibrium and therefore, on *a priori* grounds, far more probable¹².

These considerations receive an entirely new twist in the light of the so-called inflationary Universe scenario. The essential idea is that at around 10^{-35} s the Universe was in a hot phase in which the grand unified force had not yet differentiated into its nuclear and electromagnetic components. Once the temperature had dropped below the GUTs value of around 10^{30} K, the way lay open for a 'freeze out' to occur through a phase transition. One consequence of this transition is an enormous

jump in the value of the cosmological constant, Λ . Being essentially zero today, this constant must have been colossal in the early Universe and, driven by the associated repulsion, the Universe would have embarked on a period of exponential expansion with a GUTs e-folding time.

Given long enough, the inflationary phase would be capable of smoothing out any initial anisotropy and inhomogeneity by 'stretching the bumps'. At the end of inflation, as pointed out by Guth¹, the Universe would be close to the uniform structure we observe today.

We are thus led to the following scenario. In the beginning, quantum gravity dominated. At the end of the Planck era (10^{-43} s) spacetime was irregular on all scales (see, for example, fractal spacetime in ref. 7), in the form of thermal spacetime 'foam', representing maximum entropy. The Universe therefore began in an arbitrary, rather than remarkably specific, state. This is precisely what one would expect if the Universe is to be explained as a spontaneous random quantum fluctuation from nothing. Because the GUTs time is not much longer than the Planck time, the quasi-classical post-planckian era would have been too short for significant black hole formation to have occurred before inflation set in, and any holes so formed would have been microscopic and would therefore have rapidly evaporated.

Inflation began when the thermal energy from the Planck foam was redshifted below the false GUTs vacuum energy by the expansion. At this point the crucial mechanism for the gravitational winding of the Universe was set in motion. The aforementioned tendency of the gravitational field to grow clumpy with time only applies so long as the cosmological term is negligible. If the cosmological repulsion dominates, the tendency for spacetime to crumple is reversed. Instead it becomes unstable against exponential growth with asymptotic approach to (smooth and maximally symmetric) de Sitter space. de Sitter space represents the equilibrium end state (maximum entropy state) of a spacetime with a non-zero Λ .

There is thus a close parallel between the collapse of matter to form a black hole, which wipes out most of the information about the initial state (the 'no-hair' property) and the inflationary Universe, which wipes out the details of the initial cosmological phase. Both can be regarded as efficient entropy-generating (information-destroying) processes. A de Sitter 'no-hair' theorem has been conjectured by Hawking and Moss¹³. Supporting evidence for this conjecture comes from the work of Frieman and Will¹⁴ and Barrow¹⁵.

This curious two-edged instability of gravity in the presence of a cosmological constant has been known since the early analysis of the Einstein Universe. This static model is doubly unstable against collapse or exponential expansion. Note that if the initially static Universe has radius R , then according to the Einstein equations

$$R \sim GM \quad (1)$$

where M is the total mass. Equilibrium (unstable) requires

$$\Lambda \sim GM/R^3 \quad (2)$$

We may now compare the entropy of the inflationary fate with that of collapse by comparing the area of the de Sitter horizon (taken to be the gravitational entropy of de Sitter space¹⁶) with that of an equivalent-mass black hole. From equations (1) and (2) the former ($\sim \Lambda^{-1}$) is clearly $\approx$ the latter ($\sim G^2 M^2$).

So long as Λ persists, it remains entropically favourable for the Universe to continue inflating and decrinkling. However, exponentiation drops the temperature rapidly to nearly zero, and the false quantum vacuum becomes unstable against a transition to the true vacuum. Whether this takes place by supercooling and quantum tunnelling, or slow decay, is irrelevant here. What matters is that once Λ vanishes, the entropic situation is abruptly transformed again. It becomes favourable once more for the Universe to grow clumpy. Suddenly it is in a state of much-less-than-maximum entropy.

In compensation there is a huge and sudden increase in the matter entropy as the change of quantum state from an essentially zero-entropy vacuum state to a high temperature thermal state takes place. This thermal release has been called the 'grand bang' by Linde¹⁷ and is a classic example of entropy generation caused by a lag behind equilibrium conditions. The quantum state, either because of supercooling or sluggishness, hangs in the 'wrong' state for many inflationary e-folding times. During this phase the famous 'gap' opens up between the maximum possible and actual entropies of the quantum field. When it stops, a huge quantity of entropy is dumped into the matter fields. This grand-bang entropic leap is irreversible. A recontracting Universe arriving at the big crunch would not undergo 'deflation', for this would require an exceedingly improbable conspiracy of quantum coherence to reverse-tunnel through the phase transition. There is thus a distinct and fundamental asymmetry between the beginning and end of a recontracting Universe caused by the specific interplay between the quantum state and gravitational behaviour of the high temperature Universe.

Natural though this inflationary explanation of time asymmetry may be, we still seem to be left with a paradox. Although we can trace dynamically how the Universe can start out in thermodynamic equilibrium at the Planck era, but still achieve its present relatively smooth gravitational condition and less-than-maximum matter entropy, the total system of matter and gravitational field ought to be closed. There is then a thermodynamic puzzle as to how a closed system can progress from a maximum entropy to less than maximum entropy. Whereas the changing cosmological gravitational field can drive an entropy gap for the matter by acting as a changing constraint, what constraints change for the total system?

The resolution of this paradox comes from a careful consideration of the nature of energy in a cosmological context. To say that entropy is a maximum means 'a maximum for the given energy'. If the energy goes up, the entropy can go up too. Thus, a box of thermal radiation will have maximum entropy, but if more energy is introduced into the box (that is the temperature is raised) the entropy will rise. Now in cosmology, energy is not conserved. This is illustrated dramatically by the inflationary Universe scenario, where the cosmological constant acts like a fluid with a negative pressure. Thus, as the Universe expands, negative work is done, and the Λ energy in a co-moving volume rises. In fact, Λ acts such that the energy density remains constant in spite of the expansion. At the end of the inflationary phase this huge additional energy is thermalized and dumped into the matter (matter includes radiation) fields. This is the energy which provides Guth's so-called 'free lunch'¹⁸. Thus, the matter entropy, which was a maximum at the Planck era (Planck temperature at the Planck time) is now enormously greater than it would have been after inflation, had not the Λ energy thermalized.

This huge increase in matter entropy could more than offset any decrease in gravitational entropy. In the absence of a proper theory of gravitational entropy we cannot say, however, whether the inflationary stretching of the irregularities left the gravitational entropy in a co-moving volume unchanged, or actually lowered it. Nevertheless, the gravitational entropy density fell rapidly as the inflation generated more and more space.

The effect of inflation, then, was vastly to enhance the ratio of matter entropy to gravitational entropy, which at the Planck time was probably of the order of unity. We therefore have an explanation for why the gravitational entropy of the Universe is so low. The huge imbalance between the observed entropy of matter and radiation, and the gravitational entropy the Universe would have, had all this energy collapsed into a black hole, arose because the Λ energy was converted to matter entropy and not gravitational entropy at the end of the inflationary era. Since then, gravity has been trying to redress the balance. We thus have a vast reservoir of relatively negative gravitational entropy.

In conclusion, it is now possible, using the inflationary scenario, to postulate a Universe which begins in an arbitrary, equilibrium state, both as far as matter and gravity are concerned but which, because of the linkage between the quantum state of the matter and the gravitational behaviour through Λ , first gets wound up gravitationally by inflation, then gets wound up in the nuclear sense at a much later epoch by the conventional expansion. The remaining history of the Universe is the subsequent attempt to unwind by gravitational clumping (galaxies $\rightarrow$ stars $\rightarrow$ blackholes $\rightarrow$ superholes) and nucleosynthesis (hydrogen $\rightarrow$ helium $\rightarrow$ iron). Together, these two evolutionary chains account for all the observed macroscopic time asymmetry in the physical world and imprint on our environment a distinct arrow of time.

Received 26 August; accepted 26 November 1982.

1. Guth, A. H. *Phys. Rev. D* **23**, 347 (1981).
2. Linde, A. D. *Phys. Lett. B* **108**, 389 (1982).
3. Hawking, S. W. & Moss, I. G. *Phys. Lett. B* **110**, 35 (1982).
4. Albrecht, A. & Steinhardt, P. J. *Phys. Rev. Lett.* **48**, 1220 (1982).
5. Davies, P. C. W. *The Physics of Time Asymmetry* (Surrey University Press—University of California Press, 1974).
6. Frautschi, S. *Science* **217**, 593 (1982).
7. Davies, P. C. W. in *Quantum Gravity 2: A Second Oxford Symposium* (eds Isham, C. J., Penrose, R. & Sciama, D. W.) (Clarendon, Oxford, 1981).
8. Penrose, R. in *General Relativity: An Einstein Centenary Survey* (eds Hawking, S. W. & Israel, W.) (Cambridge University Press, 1979).
9. Barrow, J. D. *Phys. Rev. Lett.* **46**, 963 (1981).
10. Bekenstein, J. D. *Phys. Rev. D* **7**, 2333 (1973).
11. Hawking, S. W. *Commun. Math. Phys.* **43**, 19 (1975).
12. Barrow, J. D. *Nature* **272**, 211 (1978).
13. Hawking, S. W. & Moss, I. G. *Phys. Lett. B* **110**, 35 (1982).
14. Frieman, J. A. & Will, C. M. *Astrophys. J.* **259**, 437 (1982).
15. Barrow, J. D. in *Very Early Universe* (eds Gibbons, G., Hawking, S. W. & Siklos, S. T. C.) (Cambridge University Press, 1983).
16. Gibbons, G. W. & Hawking, S. W. *Phys. Rev. D* **15**, 273 (1977).
17. Linde, A. D. *Phys. Lett. B* **99**, 391 (1981).
18. Guth, A. H. in *Asymptotic Realms of Physics: A Festschrift in Honor of Francis Low* (eds Guth, A. H., Huang, K. & Jaffe, R. L.) (MIT Press, 1983).

Optical variability of QSOs and gravitational lenses in galactic haloes

Roland Bacon* & Jean-Luc Nieto†

Observatoires du Pic-du-Midi et de Toulouse, Laboratoire Associé au CNRS No. 285, 65200 Bagnères-de-Bigorre, France

The problem of QSO-galaxy associations has been analysed^{1,2} by taking into account different classes of QSOs with the conclusions that: (1) the strong excess of 3C QSOs near bright galaxies³ is not confirmed by larger QSO samples, which do not show any tendency towards QSO-galaxy associations, so that the 3C QSOs appeared as a statistical fluctuation¹. (2) Amongst various classes of QSOs, optically variable (OV) QSOs from a list of 41 objects⁴ showed a significant excess of QSOs near galaxies². But to be considered as a physical effect this excess needed confirmation from a larger sample of QSOs. Canizares⁵ has used the possibility of gravitational lens effects on QSOs caused by stars in galactic haloes to explain the presence of numerous QSOs near bright galaxies⁶, whose statistical significance is quite controversial (see refs 7–11). Canizares attributed to such QSOs one physical characteristic—their optical variability caused by the relative motions of the Earth and of the star acting as a gravitational lens. As these theoretical considerations gave a new insight on our previous statistical study, we decided to consider a larger sample of OV QSOs. We now present this new analysis and discuss the results in the light of Canizares' considerations.

We followed the same procedure as in ref. 2. The galaxies were taken from the Zwicky catalogues ($m \leq 15.7$ and $\delta > -3^\circ 30'$). The QSOs came from the list set up by Nieto and Seldner¹² made up of QSOs having an accurate position and a redshift determined with some confidence provided $z > 0.1$.

* Present address: Observatoire de Lyon, 69230 Saint-Genis-Laval, France.

† To whom correspondence should be addressed.

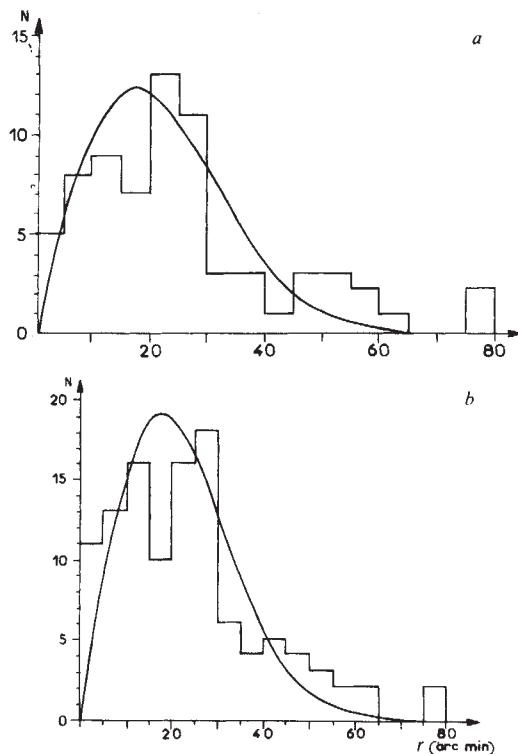


Fig. 1 Histograms of the corrected angular distances between OV QSOs and their nearest galaxy compared with the expected distribution. *a*, Sample (2); *b*, sample (3).

From this list, we extracted those objects indicated as OV in ref. 13 so that 112 QSOs were counted with $\delta > -3^{\circ}30'$.

We shall call sample (1) the subsample of 41 objects already used in ref. 2. Sample (2) will correspond to the 71 new objects and sample (3) to the whole new sample. The distances between QSOs and the first nearest galaxy were measured and corrected for the local density of galaxies in $6^{\circ} \times 6^{\circ}$ area around the QSO. The reference density of galaxies was the average density of galaxies brighter than $m = 15.7$, $\mu_0 = 1.79 \text{ (deg)}^{-2}$. Figure 1 shows the histograms of the corrected distances from samples (2) and (3) compared with the expected distribution as discussed in ref. 1. See ref. 2 for the histogram from sample (1). We clearly see an excess of small angular distances ($r < 5$ arc min) as summarized in Table 1.

It has been shown previously^{1,2}, in particular with numerical simulations, that: (1) the whole shape of such observed histograms is consistent with a random distribution of QSOs with respect to galaxies; (2) in particular, the departures from the distribution expected when assuming a local constant density around the QSO are essentially due to the effect of a small-scale clustering of galaxies in this local area; (3) at small angular distances, the small-scale clustering has no effect and cannot explain a strong departure such as the excess found here.

Sample (3) confirms the excess found with sample (1), but sample (2) fails to show significantly the same effect.

We may wonder about the reliability of the correction in a $6^{\circ} \times 6^{\circ}$ area. For instance, a correction in a $1^{\circ} \times 1^{\circ}$ area instead would put some corrected distances out of the 5 arc min bin, at least in the case of sample (1) (see ref. 2). It turns out, in fact, that such a change in the corrected distances is much smaller than a factor of 2, meaning that all the QSOs in the 5 arc min bin would still go into the 10 arc min bin. Even with such a drastic change, the statistics in the 10 arc min bin would still give significant results (see also ref. 2, Table 1).

Is this statistical difference between samples (1) and (2) related to physical properties of their QSOs? Sample (2) QSOs are on average fainter than sample (1) QSOs: $\bar{m}_2 = 17.5 \pm 0.3$ compared with $\bar{m}_1 = 16.8 \pm 0.2$. We therefore repeated the

Table 1 Characteristics of the histograms of corrected distances from the three QSO samples

Sample	<i>N</i>	Observed	Expected	<i>P</i>
1	41	6	1.6	5×10^{-3}
2	71	5	2.8	1.4×10^{-1}
3	112	11	4.4	4×10^{-3}

N is the number of objects in the sample. Columns 3 and 4 respectively show the observed and the expected numbers of angular distances with $r < 5$ arc min. *P* is the corresponding probability (Bernoulli test).

Table 2 Characteristics of the histograms from bright and faint QSOs

Sample	<i>N</i>	Observed	Expected	<i>P</i>
$m < 17.5$	51	6	2.0	1.3×10^{-2}
$m \geq 17.5$	61	5	2.4	8.4×10^{-2}

See Table 1 for specifications.

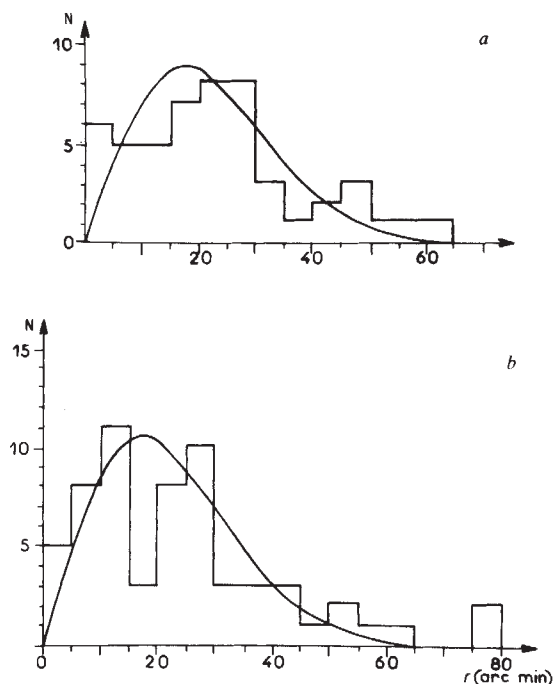


Fig. 2 Histograms of the corrected angular distances between OV QSOs and their nearest galaxy compared with the expected distribution. *a*, QSOs brighter than $m = 17.5$; *b*, QSOs fainter than $m = 17.5$.

analysis for the brightest ($m < 17.5$) and the faintest ($m > 17.5$) QSOs of our sample. Figure 2 shows the corresponding histograms. The results at small angular distances are reported in Table 2.

A tendency for brighter QSOs to present a higher excess of close associations than fainter QSOs is suggested here. Is this excess related to the apparent brightness of the objects? Clearly, more data are needed to decide on this particular point.

We have isolated a very-well defined class of QSOs, namely OV QSOs, which occur in higher numbers close to galaxies than a random distribution would predict. We also seem to see a possible dependence of such an excess with the apparent brightness of the QSOs. (There is, on the other hand, no such dependence with their absolute brightness.)

The two characteristics concerning the optical variability and the apparent brightness were predicted by Canizares⁵ when discussing the possibility of lens effects due to gravitational

focusing of the light of QSOs that are nearly co-aligned with individual stars in galactic haloes. From such a phenomenon, apparent brightness enhancement of the image of the QSO is expected as well as optical variability caused by the relative motions of the Earth and of the star acting as a gravitational lens.

Is it possible then, with our calculations, to determine the amplification factor, namely the ratio of observed QSOs to expected QSOs near galaxies, as discussed by Canizares? Our amplification ratio is of the order of 2–3 whereas Canizares presents a value hardly higher than 1. Many uncertain parameters are involved in his calculations which concern only one single star as lensing source: the maximum linear dimension of galactic haloes and the velocity dispersion of the stars in the halo assumed to be an isothermal sphere. Canizares also uses for the original distribution of QSOs (that is of unlensed QSOs) the observed density of QSOs per square degree which might be affected by gravitational amplifications (see ref. 14). On the other hand, we have considered here just OV QSOs whereas Canizares' estimate concerns all types of QSOs. In fact there must be two kinds of OV QSOs: intrinsically variable QSOs, and QSOs for which the variability is produced by gravitational lenses such as those discussed above. It is likely that this second kind of QSOs contains objects that also belong to the first kind. There is no way at present to exclude the absence of variability in a QSO, nor to discriminate between these two kinds of variability using, for instance, the signature of a typical variability due to lensing as discussed by Canizares. Then our OV samples, probably made up with these two kinds of QSOs, are not quite pure nor complete for quantitative comparison with theoretical expectations, but it is quite encouraging to see a qualitative agreement.

A distribution of QSOs at random with respect to galaxies would look non-random because the QSOs which are by chance on the line of sight of a galactic halo would be submitted to gravitational lens effects. These assumptions, supported by our statistical calculations, lead to a series of suggestions of major astrophysical importance, such as the existence of huge galactic haloes and of gravitational lens effects involving visible objects (namely the QSO and the galaxy that the lensing star belongs to) at large angular distances; much larger than the small angular distances involved in the few observational cases known^{15–18}.

QSOs would then not be at the distance of the optically associated galaxy. No well-defined sample of QSOs larger than the 3C sample had confirmed the excess found in these objects, except our samples of OV objects^{1,2}. We suggest that this excess of 3C QSOs near bright galaxies is caused by OV QSOs 'contaminating' the statistics and corresponding to a specific kind of objects distinct from the other QSOs. As suggested by Canizares³ some of Arp's objects⁶ close to galaxies might also be subject to such effects. If gravitational lenses can suggest a solution to the general problem of QSO–galaxy associations, one alternative (but not the only one) to some QSOs not being at the distance of the galaxies they are close to, is that they would be at their cosmological distances.

Received 16 August; accepted 17 November 1982.

1. Nieto, J.-L. *Astr. Astrophys.* **70**, 219–226 (1978).
2. Nieto, J.-L. *Astr. Astrophys.* **74**, 152–155 (1979).
3. Burbridge, E. M., Burbridge, G. R., Solomon, P. M. & Strittmatter, P. A. *Astrophys. J.* **170**, 233–240 (1971).
4. Grandi, S. A. & Tift, W. G. *Publ. Astr. Soc. Pacif.* **86**, 873–884 (1974).
5. Canizares, C. R. *Nature* **291**, 620–623 (1981).
6. Arp, H. C. *Astrophys. J.* **250**, 31–42 (1981).
7. Wills, D. *Phys. Scr.* **17**, 333–337 (1978).
8. Nieto, J.-L. *Nature* **270**, 411–412 (1977).
9. Burbridge, G. R. *Nature* **282**, 451–455 (1979).
10. Browne, I. W. A. *Astrophys. J. Lett.* **263**, L7–L8 (1982).
11. Webster, A. *Mon. Not. R. astr. Soc.* **200**, 47P–49P (1982).
12. Nieto, J.-L. & Seldner, M. *Astr. Astrophys.* **112**, 321–329 (1982).
13. Hewitt, A. H. & Burbridge, G. R. *Astrophys. J. Suppl.* **43**, 57–158 (1980).
14. Turner, E. L. *Astrophys. J.* **242**, L135–L139 (1980).
15. Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. *Astrophys. J.* **241**, 507–520 (1980).
16. Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. *Astrophys. J.* **244**, 736–755 (1980).
17. Weedmann, D. W., Weymann, R. J., Green, R. F. & Heckman, T. M. *Astrophys. J. Lett.* **255**, L5–L9 (1982).
18. Weymann, R. J. *et al. Nature* **285**, 641–643 (1980).

Formation of complex molecules in TMC-1

A. Freeman & T. J. Millar

Mathematics Department, UMIST, Sackville Street, PO Box 88, Manchester M60 1QD, UK

The Taurus molecular cloud 1 (TMC-1) is an interstellar gas cloud which contains a remarkable variety of complex molecules, including members of the cyanopolyne group [HC_nN , $n=3, 5, 7, 9$] and other, possibly related, species, including C_4H , C_3H_4 and C_3N . The cyanopolynes, which are found in several, different interstellar environments, have a notably high abundance in TMC-1 for such a small cloud and their method of formation has been a considerable problem in gas-phase chemistry. Several authors^{1–3} have presented schemes for the production of the cyanopolynes but all these schemes break down under the weight of observational and experimental evidence⁴. We have reinvestigated the mode of cyanopolyne formation in TMC-1 and find that their absolute abundance may be reproduced to within a factor of 3 by a model which uses gas-phase reactions of atomic nitrogen with hydrocarbons to produce them. The observed, spatial displacement of the cyanopolyne and ammonia peaks⁵ is readily understood in this model and the abundances of many other molecules are found to be in excellent agreement with observations.

TMC-1 is a small (linear size 0.6 pc)⁶, dark ($A_v > 2.4$ mag)⁷, quiescent, cold ($T \sim 10$ K), disk-shaped cloud situated in, or near, Heiles' cloud 2 in the constellation of Taurus. It has proved a particularly rich source of interstellar molecules, including the cyanopolynes, a complete list of which can be found in Table 1. The data in Table 1 represent a substantial fraction of the currently known interstellar molecules^{8,9}. The cyanopolynes have been detected in several other sources³ including the envelope of the red-giant star IRC+10216, the Orion molecular cloud and Sgr B2, the cloud almost at the centre of our Galaxy.

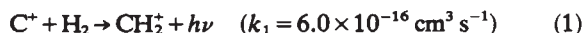
The physical conditions in TMC-1 needed to model the chemistry within the cloud may be deduced, within limits, from the observations listed in Table 1^{10,11}. The cloud density, n , is set at $3 \times 10^4 \text{ cm}^{-3}$ (the total H-nuclei number density), the temperature, T , is ~ 10 K and the fractional abundance of electrons, $x(e)$, is low at $10^{-8} \rightarrow 10^{-7}$ (ref. 12). This last parameter is very important, as it determines the dominant form of chemistry which takes place in a particular cloud source¹³. In TMC-1, the cyanopolynes and so on, appear to be spatially co-existent with molecular ions which are produced efficiently by H_3^+ chemistry only (such as, HCO^+ , N_2H^+ and their deuterated versions DCO^+ , N_2D^+). The main point about H_3^+ chemistry is that it is most efficient in low-electron, high extinction environments. The high extinction of UV photons by dust grains means that chemical processes are initiated by cosmic ray (CR) ionization (for example, $\text{H}_2 + \text{CR} \rightarrow \text{H}_2^+ + e$), as opposed to photoionization by the diffuse, UV starlight present in our Galaxy (for example, $\text{C} + h\nu \rightarrow \text{C}^+ + e$). Our model for TMC-1 includes all of the above factors. We found that depletion factors of 3 for both C and O gave best agreement with observation.

The physical model we have adopted for TMC-1 is a semi-infinite slab, with a proper treatment of radiative transfer¹⁴. To agree with the picture of an embedded, edge-on disk we used only the inner three-tenths of the model to represent TMC-1, the remaining portion only providing an extinction of 5 mag surrounding the cloud, representing material of the 'host' cloud, Heiles' 2. We then calculated the abundances of various chemical species at several points in the slab, each point providing the abundance along a 'slice' of the slab. Observers, looking at an edge-on disk, would see molecules in a column across the

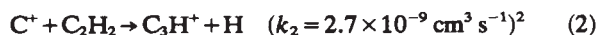
face of these slices, their column densities being $Rn(X)$, where $n(X)$ is the number density across a surface of species X and R is the disk depth.

The chemical model we have used is a simple oxygen-carbon-metal chemical reaction set^{4,15} for which an equilibrium solution was obtained, within the physical framework described above. The abundances thus calculated were then used to estimate the abundances of more complicated molecules¹⁶ (see Table 2 for the calculated fractional abundances at $A_v = 5$ mag in our model).

We propose that the cyanopolyynes and related molecules are produced by a simple set of reactions, involving the buildup of hydrocarbons through reactions of C^+ , H_2 and electrons, followed by reactions with atomic nitrogen. In the scheme, formation of complex hydrocarbons is initiated by the radiative association,



At low T and $x(e)$, further reactions with H_2 lead to the ion CH_3^+ . This ion can radiatively associate with H_2 with a rate coefficient $k = 4.5 \times 10^{-16} (T/300)^{-2.3} \text{ cm}^3 \text{ s}^{-1}$ (ref. 17) to form CH_5^+ which can recombine with an electron to form methane, CH_4 , or the methyl radical, CH_3 . As the dominant loss route for CH_3^+ is with H_2 rather than electrons, the methane abundance is independent of the radiative association rate for the $CH_3^+ - H_2$ reaction but is proportional to k_1 . At relatively high ($A_v > 3$) extinctions, methane is chiefly destroyed by carbon ions and proceeds to the acetylenic ions², $C_2H_2^+$ and $C_2H_3^+$. The latter ion quickly undergoes electron recombination to acetylene, C_2H_2 , and the next important step is the fast reaction



The next stage we propose in our scheme is crucial—a reaction between C_3H^+ and H_2 which we require to be faster than electron recombination of C_3H^+ . The former reaction may occur in two ways:



or



Provided that the rate coefficient, k , of reactions (3) or (4) is such that

$$k > 4 \times 10^{-6} x(e) \text{ cm}^3 \text{ s}^{-1}$$

then one or the other will dominate electron recombination of C_3H^+ . At our request, D. Smith's group (University of Birmingham) have recently measured the rate coefficient of reaction (4) to be $1.4 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ at 80 K (ref. 18). Further reactions of $C_3H_2^+$ with H_2 and subsequent reactions with electrons, carbon ions and H_2 may build up even more complex 'parent' ions or molecules for the cyanopolyynes. We propose that these 'parent' species then react with atomic nitrogen to produce the appropriate cyanopolyne. For example,



Atomic nitrogen reactions are extremely difficult to study and, to our knowledge, laboratory measurements on these types of reaction have not yet been performed. Detailed calculations⁸ show that a large abundance of atomic nitrogen is required to form appropriate amounts of the cyanopolyynes. According to other models¹⁹, atomic nitrogen is the dominant form for nitrogen in interstellar gas clouds up to an extinction of ~ 5 mag. As A_v increases beyond this value, chemical processes ensure a fairly rapid transition to molecular nitrogen, N_2 , and ammonia, NH_3 . We suggest that the observed spatial difference in the peak intensities of ammonia and cyanopolyne lines is caused by just such a change in extinction; that is, the $HC_n N$ species

Table 1 Observed abundances in TMC-1

Species	Column density (cm^{-2})	Fractional abundance*	Ref.
H_2	$\sim 1.0 \times 10^{22}$	0.5	12
CH	$\sim 2.0 \times 10^{14}$	$\sim 1.0 \times 10^{-8}$	27
CO^+	5.8×10^{17}	2.9×10^{-5}	28
CN	$\sim 1.0 \times 10^{13}$	$\sim 5.0 \times 10^{-10}$	3
$CS^\ddagger$	—	—	29
OH	$\sim 6.0 \times 10^{14}$	$\sim 3.0 \times 10^{-8}$	27
SO	5.0×10^{13}	2.5×10^{-9}	30
$HCN^\ddagger$	—	—	31
$HNC^\ddagger$	—	—	31
C_2H	8.4×10^{13}	4.2×10^{-9}	32
$HCO^+_\S$	—	$\sim 8.0 \times 10^{-9}$	12
$DCO^+_\S$	—	$1.04 \rightarrow 1.36 \times 10^{-10}$	33
$N_2H^+_\S$	—	—	31
$N_2D^+_\S$	—	—	31
NH_3	$\sim 1.0 \times 10^{15}$	$\sim 5.0 \times 10^{-8}$	6
C_3N	7.0×10^{12}	3.5×10^{-10}	34
H_2CO	$\sim 1.2 \times 10^{14}$	$\sim 6.0 \times 10^{-9}$	27
$HNCO$	$\sim 1.8 \times 10^{12}$	$\sim 9.0 \times 10^{-11}$	27
HC_3N	6.0×10^{13}	3.0×10^{-9}	34
DC_3N	$1.2 \rightarrow 4.8 \times 10^{12}$	$0.6 \rightarrow 2.4 \times 10^{-10}$	33
C_4H	1.6×10^{14}	8.0×10^{-9}	35
HC_5N	1.0×10^{14}	5.0×10^{-9}	6
DC_5N	$1.6 \rightarrow 6.0 \times 10^{12}$	$0.8 \rightarrow 3.0 \times 10^{-10}$	36
C_3H_4	$4.0 \rightarrow 8.0 \times 10^{13}$	$2.0 \rightarrow 4.0 \times 10^{-9}$	35
HC_7N	2.5×10^{13}	1.25×10^{-9}	6
HC_9N	3.0×10^{12}	1.5×10^{-10}	3

* Fractional abundance = $N(X)/2N(H_2)$ unless otherwise indicated.

† This CO may lie in the low-density gas surrounding TMC-1.

‡ Detections only.

§ Fractional abundances only given ($n(X)/n$).

Table 2 Calculated abundances in TMC-1 at $A_v = 5$ mag

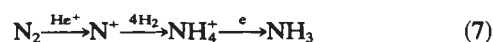
Species	Fractional abundance	Species	Fractional abundance
C	1.95×10^{-7}	HCN	2.13×10^{-9}
N	1.28×10^{-5}	CN*	1.41×10^{-10}
O	3.74×10^{-5}	C_2	3.78×10^{-11}
M (metals)	9.83×10^{-11}	C_2H^*	4.20×10^{-9}
H_2^*	0.5	C_2H_2	1.85×10^{-7}
CH^+	1.11×10^{-10}	$C_3H_2^+$	2.67×10^{-11}
OH^*	4.85×10^{-8}	$C_3H_4^*$	1.94×10^{-8}
CO^+	1.86×10^{-4}	C_3H_3	5.16×10^{-10}
H_2CO^*	1.69×10^{-8}	$CH_3CH^\ddagger$	$1.32 \times 10^{-10} \rightarrow$ 1.20×10^{-8}
CH_4	2.78×10^{-7}	HC_3N^*	2.37×10^{-9}
e	1.76×10^{-8}	H_3C_3N	3.36×10^{-10}
C^+	5.08×10^{-9}	C_4H^*	9.70×10^{-9}
$M^+ \left(\begin{smallmatrix} \text{metal} \\ \text{ions} \end{smallmatrix} \right)$	5.90×10^{-9}	C_3N^*	1.02×10^{-9}
H_3^+	5.46×10^{-10}	C_3H_3	1.00×10^{-9}
HCO^{**}	3.98×10^{-9}	HC_5N^*	5.00×10^{-9}
CH_2^+	6.98×10^{-12}	N_2H^+	1.0×10^{-10}
CH_3^+	2.05×10^{-11}	N_2	1.28×10^{-5}

* Within a factor of 3 of the observed value (Table 1).

† Not in good agreement.

‡ Value dependent on choice of destruction route.

are formed most efficiently at the edge of our TMC-1 model (at $A_v = 5$ mag) and NH_3 is formed most effectively at the centre ($A_v = 7$ mag), through the simplified scheme²⁰,



Some of the species produced in our reaction scheme have formation routes which are independent of the hydrocarbon formation mechanism (such as reactions (1), (2), (3) or (4)) that is they can be formed using known reactions involving known

Table 3 Calculated abundances in L134 N

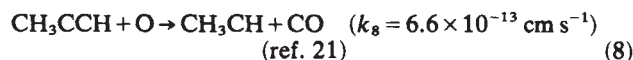
Species	Fractional abundance	Species	Fractional abundance
C	1.40×10^{-8}	HCN*	1.75×10^{-9}
N	1.28×10^{-5}	CN*	5.46×10^{-10}
O	4.25×10^{-5}	C ₂	1.15×10^{-9}
M (metals)	1.00×10^{-10}	C ₂ H*	1.37×10^{-9}
H ₂ *	0.5	C ₂ H ₂	1.99×10^{-8}
CH ⁺	6.92×10^{-11}	C ₃ H ₅ ⁺	2.71×10^{-12}
OH ⁺	3.43×10^{-7}	C ₃ H ₄ *	1.47×10^{-9}
CO*	6.17×10^{-5}	C ₃ H ₃	3.31×10^{-11}
H ₂ CO*	1.31×10^{-8}	CH ₃ CH ⁺	$3.23 \times 10^{-11} \rightarrow 5.99 \times 10^{-9}$
CH ₄	8.65×10^{-8}	HC ₃ N*	2.00×10^{-10}
e	2.31×10^{-8}	H ₃ C ₃ N	5.04×10^{-11}
C ⁺	3.44×10^{-9}	C ₄ H	1.11×10^{-9}
M ⁺ (metal ions)	5.90×10^{-9}	C ₃ N ⁺	5.44×10^{-10}
H ₃ ⁺	3.44×10^{-9}	C ₅ H ₃	4.05×10^{-12}
HCO**	6.10×10^{-9}	HC ₅ N*	2.75×10^{-10}
CH ₃ ⁺	3.20×10^{-12}	N ₂ H ⁺	1.27×10^{-9}
CH ₅ ⁺	2.02×10^{-11}	N ₂	1.28×10^{-5}

* Within a factor of 3 of the observed value.

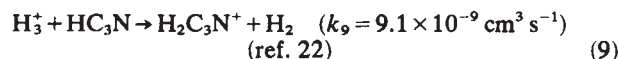
† Not in good agreement.

‡ Value dependent on choice of destruction route.

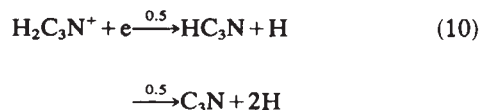
constituents of TMC-1. For example,



and



followed by



can produce CH₃CH and C₃N. The rate coefficient of reaction (8) has been measured at 300 K and could possess an activation energy. If so, it would not be important to the chemistry in TMC-1. However, if it does not possess an activation energy it becomes an efficient source of CH₃CH and may be more effective than that previously proposed by Green and Herbst²³.

We also ran a similar model for the dark cloud L134 N (L183) which has similar characteristics¹² (low T_e , $x(\text{e})$) to TMC-1 but far lower abundances of the cyanopolyynes. This second model had a lower density, a spherical geometry and $A_V(\text{centre}) = 5$ mag, consistent with the observations. In addition, we found that, to match the molecular observations, we required a greater depletion of carbon and oxygen in L134 N ($D_c = 9$, $D_o = 6$) than those we used in modelling TMC-1 ($D_c = 3$, $D_o = 3$). The results of our calculation for L134 N are shown in Table 3—again good agreement with the observations was obtained. Note that we calculate lower abundances for HC₃N and HC₅N in L134 N than in TMC-1. Two factors which account for this difference are the higher abundance of H₃⁺ and the lower abundance of CH₄ in L134 N.

The reasons why $x(\text{H}_3^+)$ is increased in L134 N when compared with TMC-1 are simple. First, because $x(\text{H}_3^+) \propto 1/n$, the lower density used in modelling L134 N results in an increase of 3 in $x(\text{H}_3^+)$. Second, both carbon and oxygen (of which CO and O are the dominant forms) are more heavily depleted. Because reactions with CO and O are the main loss routes for H₃⁺, the lower abundance of $x(\text{CO}) + x(\text{O})$ in L134 N leads to

another factor of 2 increase in $x(\text{H}_3^+)$ in this cloud. We further illustrate the differences between the chemistries in L134 N and TMC-1 by considering the species CH₄ and HC₃N.

Methane is formed by the reactions



and is destroyed mainly by C⁺, H⁺ and H₃⁺ (reactions with the latter two species tend to recycle CH₄). In TMC-1 $x(\text{H}_3^+)$ and $x(\text{H}^+)$ are less than $x(\text{C}^+)$, reaction (12) dominates the formation of CH₄, giving $x(\text{CH}_4) = 3 \times 10^{-7}$. In L134 N, since the abundance of CO is lower, less CH₃⁺ goes into CH₄. The ions H⁺ and H₃⁺ are more abundant and loss of CH₄ in reactions with them increases in importance because recycling of CH₄ by reaction (12) is less efficient. The cumulative effect is to give $x(\text{CH}_4) = 8.7 \times 10^{-8}$ in L134 N.

In TMC-1, HC₃N is formed by the reaction of C₃H₃ with N and lost by reaction with C⁺ leading to $x(\text{HC}_3\text{N}) = 2.4 \times 10^{-9}$. In L134 N, however, loss of HC₃N also occurs efficiently through reaction (9) and because 50% of the H₂C₃N⁺ so formed recombines to C₃N (reaction (10)), it is irretrievable as HC₃N. Inclusion of this loss route and the decrease in the C₃H₃ abundance (due ultimately to the reduced CH₄ abundance mentioned above) leads to $x(\text{HC}_3\text{N}) = 2 \times 10^{-10}$, as observed.

Finally, note that the displacement of the NH₃ and HC_nN peaks mentioned above may occur also in L134 N. In this cloud the cyanopolyynes have been searched for but not detected²⁵ at the position of maximum ammonia emission, whereas they have been detected at positions away from this peak²⁶. However, in some cases, line of sight orientations may cause variations from this pattern.

We thank Dr D. Smith and his group at the Space Research Department, Birmingham University for using their SIFT apparatus to measure the reaction rate of reaction (4) at such short notice. T.J.M. is supported by an SERC Advanced Fellowship.

Received 26 August; accepted 21 November 1982.

- Mitchell, G. F., Huntress, W. T. Jr & Prasad, S. S. *Astrophys. J.* **233**, 102–108 (1979).
- Schiff, H. I. & Bohme, D. K. *Astrophys. J.* **232**, 740–746 (1979).
- Winnemisser, G. & Walmsley, C. M. *Astrophys. Space Sci.* **65**, 83–93 (1979).
- Freeman, A. thesis, Univ. Manchester (1982).
- Little, L. T., Riley, P. W., Macdonald, G. H. & Matheson, D. N., *Mon. Not. R. astr. Soc.* **183**, 805–811, (1978).
- Tölle, F., Ungerechts, H., Walmsley, C. M., Winnemisser, G. & Churchwell, E. *Astr. Astrophys.* **95**, 143–155 (1981).
- Sherwood, W. A. & Wilson, T. L. *Astr. Astrophys.* **101**, 72–78 (1979).
- Blitz, L. *Scient. Am.* **246**, 72–80 (1982).
- Mann, A. P. C. & Williams, D. A. *Nature* **283**, 721–725 (1980).
- Evans, N. J. *IAU Symp.* No. 87, 1–19 (1980).
- Snell, R. L. *Astrophys. J. Suppl.* **45**, 121–175 (1981).
- Guélin, M., Langer, W. D. & Wilson, R. W. *Astr. Astrophys.* **107**, 107–119 (1982).
- Watson, W. D. *Ann. Rev. Astr. Astrophys.* **16**, 585–615 (1978).
- Black, J. H. & Dalgarno, A. *Astrophys. J. Suppl.* **34**, 405–423 (1976).
- Langer, W. D. *Astrophys. J.* **206**, 699–712 (1976).
- Huntress, W. T. Jr & Mitchell, G. F. *Astrophys. J.* **231**, 456–467 (1979).
- Adams, N. G. & Smith, D. *Chem. Phys. Lett.* **79**, 563–567 (1981).
- Herbst, E., Adams, N. G. & Smith, D. *Astrophys. J.* (submitted).
- Mitchell, G. F., Ginsburg, J. L. & Kuntz, P. J. *Astrophys. J. Suppl.* **38**, 39–68 (1978).
- Smith, D. & Adams, N. G. *Mon. Not. R. astr. Soc.* **197**, 377–384 (1982).
- Brown, J. M. & Thrush, B. A. *Trans. Faraday Soc.* **63**, 630–642 (1967).
- Freeman, C. G., Harland, P. W. & McEwen, M. J. *Astrophys. Lett.* **19**, 133–135 (1978).
- Green, S. & Herbst, E. *Astrophys. J.* **229**, 121–131 (1979).
- Prasad, S. S. & Huntress, W. T. Jr *Astrophys. J. Suppl.* **43**, 1–35 (1980).
- Snell, R. L., Schloerb, F. P., Young, J. S., Hjalmarson, A. & Friberg, P. *Astrophys. J.* **244**, 45–53 (1981).
- Benson, P. J. & Myers, P. C. *Astrophys. J. Lett.* **242**, L87–L91 (1980).
- Brown, R. L. *Astrophys. J. Lett.* **248**, L119–L122 (1981).
- Allen, M. & Knapp, G. R. *Astrophys. J.* **225**, 843–856 (1978).
- Linke, R. A. & Goldsmith, P. F. *Astrophys. J.* **235**, 437–451 (1980).
- Rydbeck, O. E. H. *et al. Astrophys. J. Lett.* **235**, L171–L175 (1980).
- Turner, B. E. & Zuckerman, B. *Astrophys. J. Lett.* **225**, L75–L79 (1978).
- Wooten, A., Bozayan, E. P., Garrett, D. B., Loren, R. B. & Snell, R. L. *Astrophys. J.* **239**, 844–854 (1980).
- Langer, W. D., Schloerb, F. P., Snell, R. L. & Young, J. S. *Astrophys. J. Lett.* **239**, L125–L128 (1980).
- Friberg, P., Hjalmarson, A. & Irvine, W. M. *Astrophys. J. Lett.* **241**, L99–L103 (1980).
- Irvine, W. M., Höglund, B., Friberg, P., Askne, P. & Eldred, J. *Astrophys. J. Lett.* **248**, L113–L117 (1981).
- MacLeod, J. M., Avery, L. W. & Broten, N. W. *Astrophys. J. Lett.* **251**, L33–L36 (1981).

Ice in Comet Bowell

H. Campins*, G. H. Rieke*† & M. J. Lebofsky†

* Lunar and Planetary Laboratory and † Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

It has been widely accepted that frozen volatiles are the major constituent of a comet nucleus¹. However, the direct detection of these ices has proved to be difficult: bright comets which are easily observable are generally so close to the Sun that icy grains are too short lived to make an appreciable contribution to the coma brightness, while comets which are far enough from the Sun for ices to survive (heliocentric distance >2 AU) are usually too faint to be observed adequately. Observations of the reflected light from comets in the 1–5 μm region of the spectrum are diagnostic of the presence of ices. Several attempts have been made to detect absorption bands in this region^{2–5}. A weak, unidentified band near 2.2 μm may have been detected in Comets Bowell and Panther, both of which are at large heliocentric distance⁵; although for Comet Bowell the result could not be confirmed⁴. We now report the detection of a deep absorption at 3.25 μm in Comet Bowell which provides the first direct evidence for the presence of H_2O ice in a comet.

The detection is the result of three nights of observation with the University of Arizona 1.54-m telescope and one night with the NASA IR Telescope Facility (IRTF) in Hawaii, using a filter centred at 3.25 μm with a bandpass of 0.45 μm . Standard K photometry (2.2 μm) was interlaced with the longer-wavelength measurements to define the shorter wavelength continuum. The comet was observed on 23 April, 11 June, and 13 June UT from Arizona and on 22 June UT (all 1982) from Hawaii. Through this period the comet was between 3.4 and 3.5 AU from the Sun and 2.9 and 2.6 AU from the Earth. 58 Ophiuchus, a solar type star (F7 IV) was used as a nearby comparison to determine the slope of the solar spectrum. The comet was measured through a 11.5 arc s diameter aperture with reference beams 14 arc s east and west of the nucleus from Arizona and with a 7.8 arc s aperture and reference beams 9 arc s to the north and south from Hawaii. All four measurement sets show the albedo at 3.25 μm to be significantly lower than that at 2.2 μm ; the combined result is that the albedo at the longer wavelength is 0.52 ± 0.06 times that at the shorter one.

This measurement is plotted in Fig. 1 along with our previous spectrophotometry near 2 μm (ref. 4); all of the measurements are normalized to the brightness at 2.2 μm . It is expected that the contaminants in the ice particles (smaller embedded dark dust grains) will bring the geometric albedo of the grains down to $\sim 5\%$ and will suppress the 1.5 and 2.0 μm ice bands, but that the stronger 3 μm band could still be present^{4,6}. As an illustration of this possibility, we show the spectrum of Callisto, which has absorption features thought to arise primarily from water frost (L. A. Lebofsky, personal communication). Note that the low surface reflectivity of Callisto has substantially decreased the relative strength of the 2.0 μm frost band in the spectrum, but a well defined band remains between 3 and 4 μm .

Our measurements are consistent with those of Comet Bowell by M. F. A'Hearn, A. Tokunaga and E. Dwek (personal communication) made with the IRTF on 27–30 April 1982. They found the albedo at 3.45 μm (bandpass 1.05 μm) to be depressed relative to the continuum at H (1.6 μm) by ~ 20 –40% and the albedo at 3.8 μm (bandpass 0.67 μm) to be depressed by ~ 15 –45%.

Of the ices which are likely to be found in comets, H_2O , NH_3 , H_2S , and CH_4 have absorption features near 3 μm (ref. 7); however, only H_2O ice is a plausible identification at the heliocentric distance of Comet Bowell when we observed it

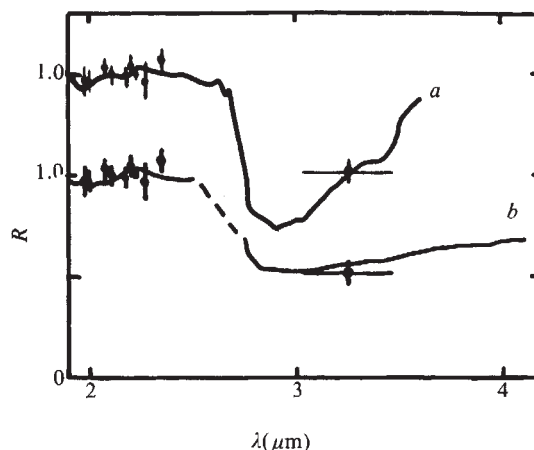


Fig. 1 Relative reflectivities (R) of Orgueil (a), Callisto (b) and Comet Bowell (points). For clarity, the spectrum of Orgueil has been displaced upward by 0.5 in R , and a duplicate set of observational points has been plotted with the same displacement. All three reflectivity curves have been normalized to 1.0 at 2.2 μm . Both hydrated minerals (curve a) and dirty ice (curve b) provide a satisfactory match to the absorption in the spectrum of Comet Bowell; from the typical gas-to-dust ratios in other comets, dirty ice is by far the more plausible fit.

(3.4 AU). From dynamical considerations, it is believed that the grains in Comet Bowell are ≥ 0.5 mm in diameter⁸. The grains have a geometric albedo of only about 6%^{5,9}, which will make their temperature approximately equal to that of a rapidly rotating black body in equilibrium with the absorbed solar radiation; at the heliocentric distance of Comet Bowell during our measurements, this temperature is about 150 K. In these conditions, it can be shown¹⁰ that the lifetimes against evaporation are respectively 1.4 yr and 12 s for H_2O and NH_3 ice grains; the other candidate ices are even more volatile than NH_3 .

These lifetimes can be compared with the time it would take a particle to travel from the nucleus of the comet to the radius included in our largest aperture (2×10^4 km). The grain velocity with respect to the nucleus is thought to be $\sim 1 \text{ m s}^{-1}$ (ref. 8), which means that the oldest grains in our beam are ~ 0.7 yr old. Therefore, H_2O grains are stable enough to have produced the observed absorption, but all of the other candidate ices are too volatile by orders of magnitude.

Another possibility is that the absorption near 3 μm is due to water of hydration in the cometary grains, as is observed in some asteroids¹¹. As an example, in Fig. 1 we have plotted the spectrum of the meteorite Orgueil¹², a CI carbonaceous chondrite with $\sim 20\%$ water (by weight) in the form of hydrated minerals. Although the bands due to water of hydration are usually narrower than those due to ice, our data alone cannot distinguish between these two cases. However, as mentioned above the data of A'Hearn *et al.* indicate that the albedo of the Comet Bowell grains is slightly reduced at 3.8 μm relative to the shorter wavelengths, in support of our interpretation that the band is due to water frost. In addition, the ratio of mass loss of H_2O gas to mass loss of solids (that is, gas-to-dust ratio) for five comets at small heliocentric distances has been estimated to range from about 1.6 to 10 (ref. 13). This ratio is therefore much larger than is typical for hydrated minerals; as it probably reflects the composition of the grains before they lose volatiles, we conclude that the grains at large heliocentric distance are composed mostly of frozen water.

The detection of ice in cometary grains provides one of the strongest possible confirmations of Whipple's¹ icy conglomerate model of cometary nuclei. It also supports a broad range of theoretical arguments that water should be the dominant parent molecule for some of the gases detected by radio and optical spectroscopy of comets. Two recent illustrations are: (1) Spinrad¹⁴ has derived the oxygen production rates for nine comets from observations of $[\text{O I}]$ lines. These production rates

are consistent with H_2O being the sole parent molecule and are in agreement with the H_2O production rates derived from IUE observations; and (2) OH has been detected in emission in Comet Bowell, both through groundbased observations and with IUE¹⁵. The OH coma of comets is thought to be a photodissociation product of H_2O (ref. 16). We thank M. F. A'Hearn for communication of unpublished observations and L. A. Lebofsky for permission to use the spectrum of Callisto in Fig. 1. Part of this research was carried out at the IRTF which is operated by the University of Hawaii under contract to NASA. Assistance in obtaining the observations with the IRTF was provided by J. Hamilton. This research was supported by NASA and by the NSF.

Received 19 July; accepted 17 November 1982.

1. Whipple, F. L. *Astrophys. J.* **111**, 375–394 (1950).
2. Oishi, M. *et al. Publ. astr. Soc. Jap.* **30**, 149–159 (1978).
3. A'Hearn, M. F., Dwek, E. & Tokunaga, A. T. *Astrophys. J. Lett.* **248**, L147–L151 (1981).
4. Campins, H., Rieke, G. H., Lebofsky, M. J. & Lebofsky, L. A. *Astr. J.* (in the press).
5. Jewitt, D. C., Soifer, B. T., Neugebauer, G., Danielson, G. E. & Matthews, K. *Icarus* (submitted).
6. Hanner, M. S. *Icarus* **47**, 342–350 (1981).
7. Fink, U. & Sill, G. T. *IAU Colloq.* **61**, 164–202 (1982).
8. Sekanina, Z. *Astr. J.* **87**, 161–169 (1982).
9. Hanner, M. S., Veeder, G. J. & Matson, D. L. *Bull. Am. astr. Soc.* **3**, 705–706 (1981).
10. Lebofsky, L. A. *Icarus* **25**, 205–217 (1975).
11. Lebofsky, L. A., Feierberg, M. A., Tokunaga, A. T., Larson, H. P. & Johnson, J. R. *Icarus* **45**, 453–459 (1981).
12. Larson, H. P., Feierberg, M. A., Fink, U. & Smith, H. A. *Icarus* **39**, 257–271 (1979).
13. Ney, E. P. *IAU Colloq.* **61**, 323–340 (1982).
14. Spinrad, H. *Publ. astr. Soc. Pacif.* (submitted).
15. Feldman, P. D., A'Hearn, M. F. & Millis, R. L. *IAU Circ. No.* 3699 (1982).
16. Keller, H. U. & Thomas, G. E. *Astr. Astrophys.* **39**, 7–19 (1975).

Stratospheric warming following the El Chichón volcanic eruption

D. E. Parker & J. L. Brownscombe

Meteorological Office, Bracknell, Berkshire RG12 2SZ, UK

The eruption of the Mexican volcano El Chichón ($17^{\circ}24' \text{N}$, $93^{\circ}00' \text{W}$) on 29 March 1982 provided the opportunity for monitoring the effects of a major volcanic eruption on the Earth's atmosphere, and thus for furthering our understanding of the causes of climatic fluctuations. The eruption injected a large quantity of material high into the stratosphere. There have been reports of dust above 30 km (A. Strong, personal communication). We now report a warming of the equatorial lower stratosphere after the eruption of El Chichón, and conclude that this warming was a combined effect of the eruption and of the tropical stratospheric quasi-biennial oscillation.

The effects of volcanic dust on stratospheric temperature have been widely studied. Newell¹ attributed a warming of about 6 K at 60–80 mbar (20–18 km) over tropical Australasia to the 1963 Mount Agung (8°S , 115°E) eruption. However, McInturff *et al.*² found that at 50 mbar (20–21 km) the tropics as a whole warmed by no more than 2 K after that eruption and they concluded that it would be difficult to isolate the volcanic effect from other influences or trends. A later study by Angell and Korshover³ indicates a 3 K warming in the tropical 100–30 mbar (16–24 km) layer in 1963 and a subsequent 2 K cooling, but they found no extratropical warming and stressed that it was not clear to what extent the tropical warming was due to Agung and to what extent it could be associated with an enhanced tropical stratospheric quasi-biennial oscillation. The temperature maximum associated with the quasi-biennial oscillation normally slightly precedes the maximum westerly wind⁴, and a warming at 50 mbar would therefore have occurred in 1963 in any case; also, because the amplitude of the quasi-biennial oscillation is not constant, it is difficult to separate volcanic and quasi-biennial effects. Thus, although Newell *et al.*⁵ tentatively ascribed the descent of the

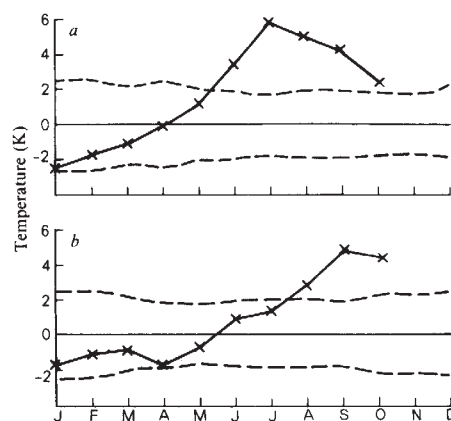


Fig. 1 Changes of temperature in the tropical lower stratosphere, 1982. *a*, 30 mbar; *b*, 50 mbar. Values are deviations from station normals averaged over Singapore (1°N 104°E), Ponape (7°N 158°E), Rieke, G. H., Lebofsky, M. J. & Lebofsky, L. A. *Astr. J.* (in the press). Dashed lines are standard deviations of monthly normals averaged over the four stations, and will not overestimate the standard deviation of the four-station normal because of the strong spatial coherence of tropical stratospheric temperature fluctuations associated with the quasi-biennial oscillation.

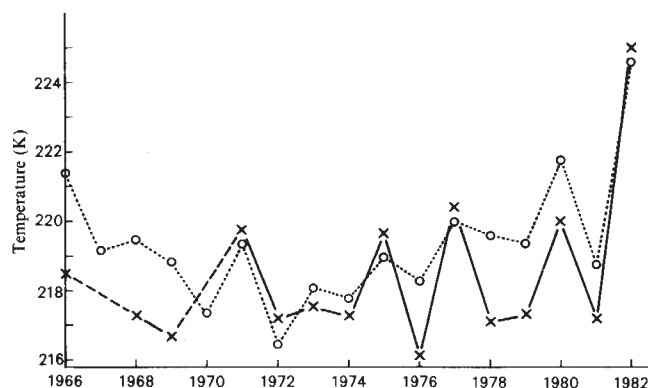


Fig. 2 Temperatures at 30 mbar in July 1966–82. $\times$, Singapore; $\circ$, Ponape.

region of anomalous warmth following Agung to settling of the dust, an alternative explanation is the normal descent⁴ of the thermal pattern of the quasi-biennial oscillation.

A study by Pollack *et al.*⁶ of radiative transfer suggests that stratospheric warming will follow injection of volcanic material into the stratosphere, mainly because of absorption of terrestrial radiation, but Newell^{1,7} considered absorption of solar radiation to be more important. Ogren *et al.*⁸ measured a significant absorption of visible radiation by stratospheric aerosol resulting from the 1980 Mount St Helens eruption.

We have used both radiosonde data and satellite radiance-based temperature data to assess the warming of the stratosphere following the eruption of El Chichón. The satellite data come from the Stratospheric Sounding Unit (SSU) and the Microwave Sounding Unit (MSU) on board the polar-orbiting NOAA 6 (see refs 9, 10 for details). The satellite data provide global coverage but their vertical resolution is of the order of 15 km. The instrumental drift of the SSU did not exceed 0.2 K between 1979 and 1982¹¹. Figures for MSU drift are not available but the reliability of the MSU is supported by its agreement with TIROS N and NOAA7 data.

Figure 1 shows deviations from normal of the near-equatorial 30 mbar ($\sim 24 \text{ km}$) and 50 mbar ($\sim 20 \text{ km}$) temperatures during 1982, based on the average of four well-spaced radiosonde stations. The station normals were the average of all available data up to 1980, that is $\sim 20 \text{ yr}$ for Ponape and $>10 \text{ yr}$ for the other stations. The greatest warmth, relative to normal, of 6 K

at 30 mbar in July was 3 s.d. from normal; and that of nearly 5 K at 50 mbar in September was 2.5 s.d. from normal. At 30 mbar the individual stations showed similar patterns, but the magnitude of the warmth was not uniform and decreased from near 8 K at the Equator to near 4 K at 9° N. At 50 mbar the station at 9° N (Howard in the Panama Canal Zone) showed no warming, Ponape (7° N) only 4 K, and the other stations 6–8 K. Figure 1 shows that the warming occurred later at 50 mbar than at 30 mbar, suggesting downward propagation, as occurred after Agung⁵. For the 100–30 mbar layer as a whole there had been a warming averaging nearly 4 K by September over these four stations, in comparison with a 3 K peak tropical warming for this layer after Agung³. The unusual warmth of July 1982 at 30 mbar over the Equator is indicated by the July temperatures for two of the stations for 1966–82 in Fig. 2. However, at both these stations August 1982 was only between 1 K and 2 K warmer than August 1977.

Figure 3 presents mean temperatures for the zone 15° N to 15° S and for the globe for 1980, 1981 and 1982. These are derived from MSU channel 24 and SSU channels X, 25 and 27, centred near 80, 50, 15 and 2 mbar respectively. Channel X is synthesized from a linear combination of the radiance of SSU channel 26 and the rate of change of radiance with angle to nadir of channel 25. This results in a rather narrow weighting function with peak response near 50 mbar. Tropical warming occurs near 50 mbar in 1982 relative to the two previous years (Fig. 3a), peaking at about 2 K in July and August. At 15 mbar (about 28 km) an earlier tropical warming is suggested, peaking in May, and at 2 mbar (about 43 km) tropical cooling is apparent in July and August, and cooling is also indicated at 5 mbar by SSU channel 26 (not shown). There is little of significance to report for 80 mbar in the tropics. The global mean temperatures (Fig. 3b), which have an r.m.s. scatter of only about 0.03 K from day to day, show significant warmings at 80, 50 and 15 mbar in 1982 relative to 1980 and 1981 but this possibly started at 80 and 50 mbar before the eruption. Examination of the latitudinal dependence of the temperature changes revealed an approximately equatorially-symmetric anomalous warming at 50 and 15 mbar, sufficient to account for the global warmth after the eruption, with a corresponding pattern of coldness at 2 mbar. The 80 mbar global warming, which was about 0.7 K, was not a mainly equatorial effect. Figure 4, based on 70 Northern Hemisphere radiosonde stations, confirms that the warming at 30 mbar had not affected middle latitudes of the Northern Hemisphere by August 1982. The stratospheric warming following the eruption of Agung was also confined to the tropics (30° N–30° S)³.

There are two reasons for caution in ascribing the warming to the eruption of El Chichón. First, the dust veil has been reported as covering the region 60° N to 20° S (A. Strong, personal communication) and because of its origin there is probably more dust in the Northern than in the Southern Hemisphere. A difficulty therefore arises in attributing the stratospheric warming to local absorption of radiation, because the warming has been approximately symmetrical about the Equator: in addition, the insolation available for absorption in April to August must have been greater in the Northern Hemisphere. Second, in the period studied the phase of the quasi-biennial oscillation has favoured downward propagating, equatorially-symmetrical¹² warmth⁴ at the levels concerned. The cooling at 2 and 5 mbar since 1981 may also be a manifestation of the quasi-biennial oscillation, but could indicate modifications to the stratospheric circulation or an alteration of the radiation balance at 2–5 mbar by absorption of terrestrial radiation by the volcanic material at lower levels. Data presented by Coy¹³ indicate that the vertical wavelength of the quasi-biennial oscillation is consistent with the first explanation, given its August 1982 phase at 30 mbar.

Nevertheless there is strong support for linking the warmth to the volcanic dust. The radiative transfer study of Pollack *et al.*⁶ provides *a priori* physical reasons for the warming. Figure 2 shows that the 1982 warmth considerably exceeded that to

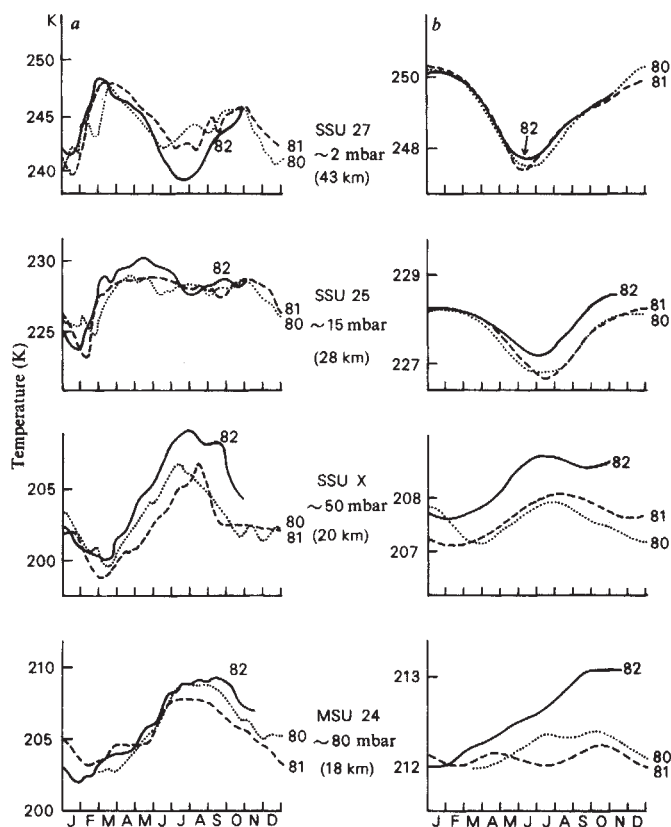


Fig. 3 Tropical (a) and global mean temperatures (b), 1980–82. Values are based on temperatures derived from NOAA 6 Stratospheric Sounding Unit and Microwave Sounding Unit radiance channels which have vertical resolution of ~15 km.

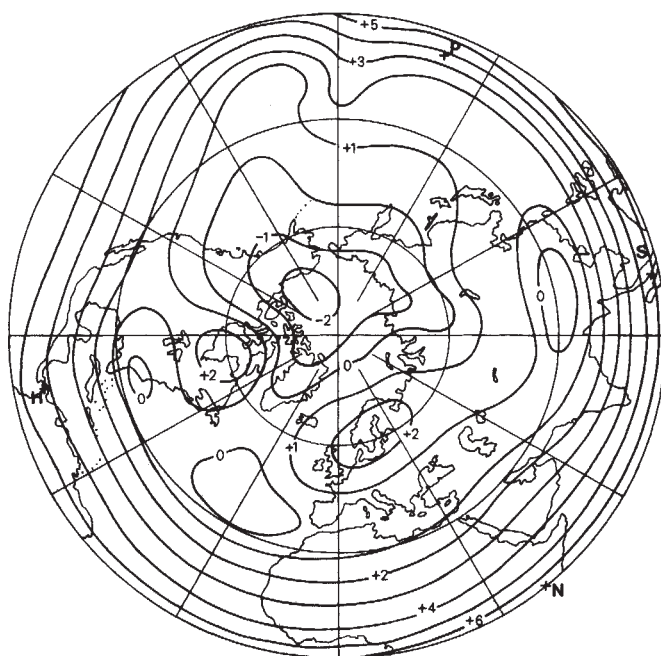


Fig. 4 30 mbar temperature differences (K) from normal, August 1982. H, Howard Air Force Base; N, Nairobi; P, Ponape; S, Singapore.

be expected from the quasi-biennial oscillation. The warming probably commenced higher in the atmosphere than after Agung, as might be expected from the known greater height reached by the El Chichon material. Of further significance is a modelling study by Holton¹² in which he demonstrates that equatorial Kelvin and Rossby–gravity waves, which are

believed^{14,15} to sustain the quasi-biennial oscillation, tend to force equatorial symmetry in the zonal flow even in the presence of initial wind fields which are anti-symmetric about the Equator, and equatorial symmetry in the resulting final temperature fields would be a corollary. Thus it can be suggested that the equatorially-symmetrical nature of the observed warming pattern was a result of the interaction between the equatorial waves which maintain the quasi-biennial oscillation and the equatorially-asymmetrical forcing of the wind fields by the thermal effect of the volcanic material; *in situ* warming by radiative absorption may have been less important at the Equator. Separation of the two effects will require not only further studies of stratospheric dynamics but also acquisition of more detailed information on the spatial distribution and radiative properties of the volcanic material. The elevated global mean temperatures deduced from the satellite data suggest anomalous absorption of heat in addition to dynamical re-distribution of heat, but a longer time series of data is required before these changes can be positively related to specific events such as the El Chichon eruption.

Dr J. Nash developed the technique for synthesizing additional SSU weighting functions and provided some of the data used in Fig. 3.

Received 28 October; accepted 7 December 1982.

1. Newell, R. E. *J. atmos. Sci.* **27**, 977-978 (1970).
2. McInturff, R. M., Miller, A. J., Angell, J. K. & Korshover, J. *J. atmos. Sci.* **28**, 1304-1307 (1971).
3. Angell, J. K. & Korshover, J. *Mon. Weath. Rev.* **106**, 1422-1432 (1978).
4. Ebdon, R. A. *Met. Mag.* **100**, 84-90 (1971).
5. Newell, R. E., Kidson, J. W., Vincent, D. G. & Boer, G. J. *The General Circulation of the Tropical Atmosphere and Interactions with Extratropical Latitudes Vol 2* (Massachusetts Institute of Technology 1972).
6. Pollack, J. B. *et al. J. geophys. Res.* **81**, 1071-1083 (1976).
7. Newell, R. E. *Nature* **227**, 697-699 (1970).
8. Ogren, J. A., Charlson, R. J., Radke, L. F. & Domonkos, S. K. *Science* **211**, 834-836 (1981).
9. Schwalb, A. *NOAA Tech. Mem. NESS 95* (1979).
10. Miller, D. E., Browncombe, J. L., Carruthers, G. P., Pick, D. R. & Stewart, K. H. *Phil. Trans. R. Soc. A296*, 65-71 (1980).
11. Nash, J. & Browncombe, J. L. *Adv. Space Sci.* (in the press).
12. Holton, J. R. *J. atmos. Sci.* **36**, 1030-1040 (1979).
13. Coy, L. *J. atmos. Sci.* **36**, 174-176 (1979) [corrigendum **37**, 912-913 (1980)].
14. Holton, J. R. & Lindzen, R. S. *J. atmos. Sci.* **19**, 1076-1080 (1972).
15. Plumb, R. A. *J. atmos. Sci.* **34**, 1847-1858 (1977).

Absolute measurement by satellite altimetry of dynamic topography of the Pacific Ocean

Chang-Kou Tai* & Carl Wunsch*†

* Department of Earth and Planetary Sciences and † Department of Meteorology and Physical Oceanography, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Satellites are the only foreseeable possibility for obtaining observations of the ocean circulation on the basin and global scales required to understand it. Such observations are essential for an understanding of the role of the ocean in various important problems (for example, climate). The 3-month Seasat mission¹ has demonstrated^{2,3} that altimetry is capable of providing global observations of oceanic variability. We show here that the data from this short, sub-optimum, mission are also adequate for a determination of the absolute sea-surface topography of the ocean on large scales—a much more difficult⁴ and even more valuable measurement. An absolute determination of the subtropical gyre of the North Pacific Ocean has been obtained; we believe this to be the first direct measurement showing the existence of such a feature, that does not depend on conventional hydrography and a series of assumptions.

The sea surface, $S(\theta, \lambda, t)$ relative to an ellipsoidal reference Earth can be written

$$S(\theta, \lambda, t) = N(\theta, \lambda) + \zeta(\theta, \lambda, t) + r(\theta, \lambda, t) \quad (1)$$

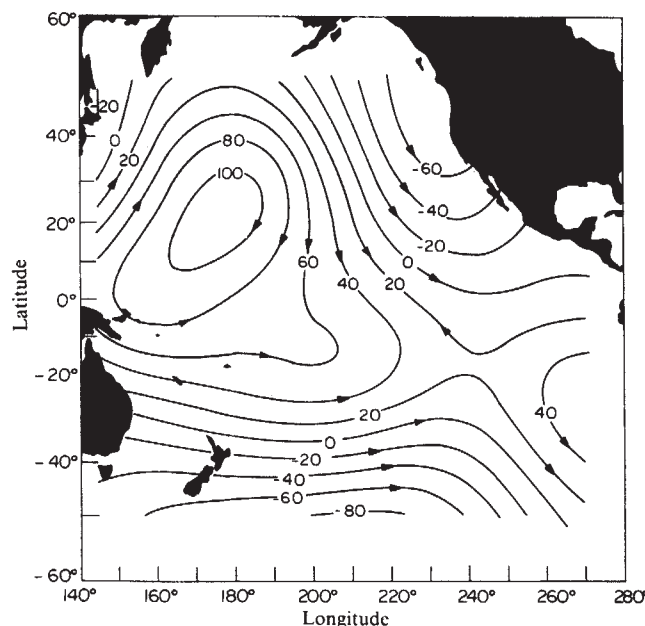


Fig. 1 The sea surface topography (cm) of the Pacific Ocean relative to the geoid as determined solely from satellite measurements. It is viewed through a low pass wavenumber filter. The arrows denote the direction of the geostrophic surface velocity which corresponds to this elevation.

where N is a particular equipotential surface of the combined gravitational and centrifugal force fields of the Earth (the geoid), ζ is the departure of the ocean from the equipotential surface owing to the large-scale movement of water, r is 'noise' that includes a variety of observational errors and geophysical phenomena (atmospheric load, for example)⁵ and θ, λ are latitude and longitude, respectively, and t is time.

An altimetric satellite measures S . The chief residual errors, after correction for various effects, are due to uncertainties in the radius of the spacecraft orbit. These errors occur primarily on the scale of the orbit diameter and manifest themselves principally through a constant offset and a tilt of the orbit. They can be drastically reduced through a crossing-arc analysis^{6,7}, that is, by requiring that successive altimetric passes which cross the same oceanic point should agree with each other within rather small limits. R. H. Rapp (personal communication) has performed such an analysis for the global Seasat data set. The r.m.s. residual errors at the crossing points are 28 cm globally and 34 cm for the Pacific alone, and are probably close to white in the spatial spectrum. Thus, we have an estimate of the 3-month average of S .

The structure of S is dominated by N (variations of order 200 m vertically relative to the reference ellipsoid). If N is sufficiently well known, one could determine ζ , the quantity which interests us from equation (1), and from the measurement of S . However, the undulations of ζ are only of order 1-2 m globally.

The so-called GEM 9 Earth gravity model⁸ represents the geoid in a spherical harmonic expansion complete to degree and order 20 (~2,000 km wavelengths) with selected higher-degree terms. The formal uncertainty is 1.9 m r.m.s. but, with the missing expansion coefficients, is estimated to have a further amplitude of about 2.5 m r.m.s. With ζ having an amplitude of only 1-2 m these errors might seem to preclude the subtraction

$$\tilde{\zeta} = S - N \quad (2)$$

necessary to obtain a useful estimate $\tilde{\zeta}$ of the 3-month average of ζ . However, GEM 9 was determined by tracking satellites in Earth orbit over the past 20 yr. The method is much more accurate in determining the coefficients of low degree and order spherical harmonics in N than high degrees. Thus, the formal

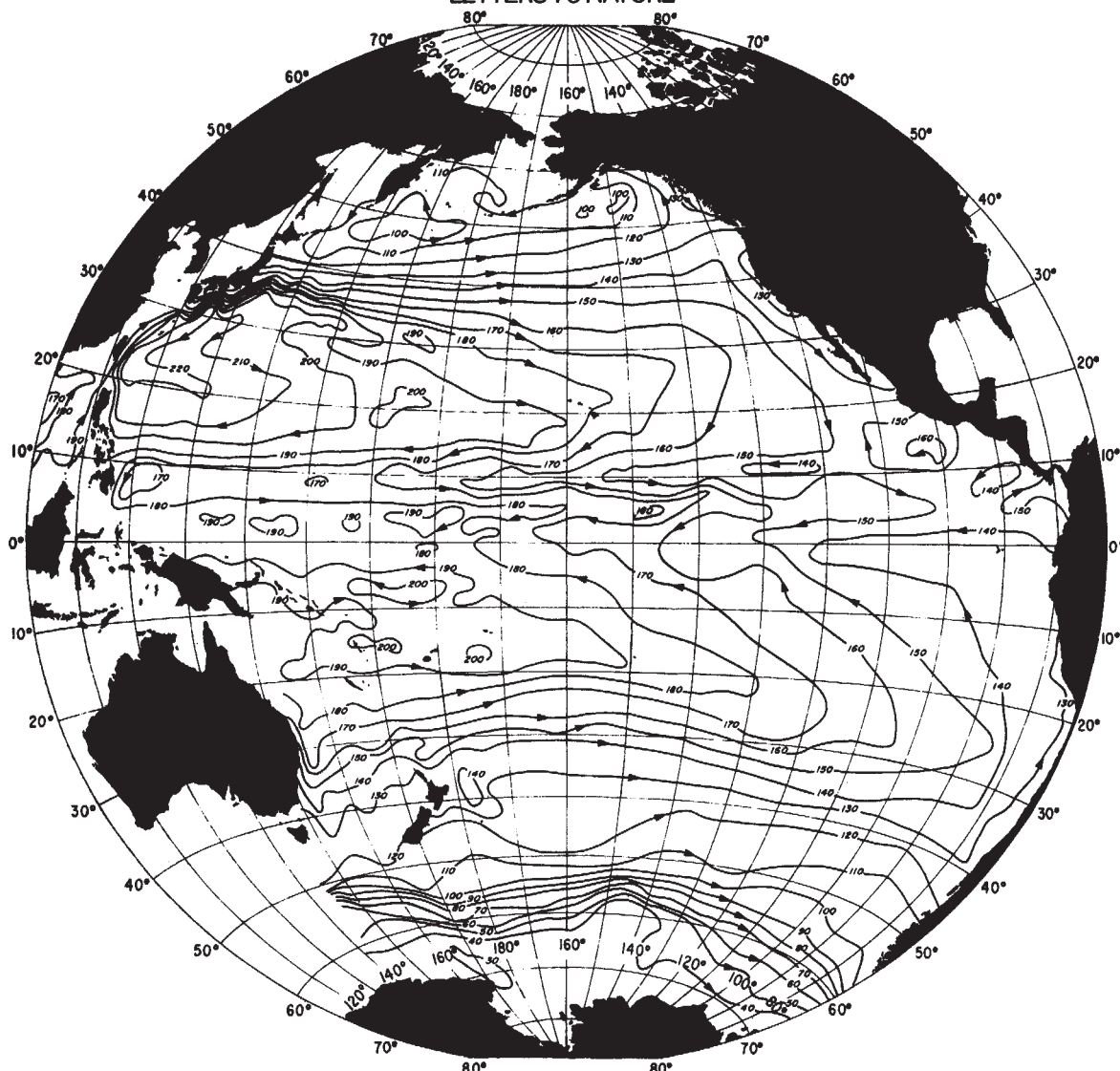


Fig. 2 Estimate¹⁰ of the sea-surface topography (dynamic cm, numerically nearly cm) of the Pacific Ocean from shipborne hydrography compiled over many years. High wavenumber components that have been removed from Fig. 1 are present in this figure. There is a strong visual correlation between the two figures.

uncertainty in GEM 9 complete to degree and order 4 (wavelengths of order 10,000 km) is about 16 cm, whereas to degree and order 6 it is about 30 cm. If the ocean circulation, as manifested through ζ , contains a significant fraction of its energy on long wavelengths, one should be able to see it with the existing estimates of N and S .

Our procedure, which is described fully by Tai⁹, is basically as follows: we subtract the GEM 9 value of N from S over the region of the Pacific where S was available to us (72° N to 72° S, 145° E to 87° W). The areal mean was removed and a 20° × 20° running average was made to remove much of the high wavenumber energy. The result was then expanded in terms of spherical harmonics. Tai⁹ discusses the resolution of the expansion from the limited area fit. We display the final form, with the expansion carried to degree and order 6, in Fig. 1, and for comparison, show in Fig. 2 Wyrki's¹⁰ estimate of the Pacific dynamic topography as computed from conventional hydrographic data (the Wyrki chart includes much high wavenumber energy excluded from Fig. 1).

Figure 3 is Wyrki's chart run through the same low-pass filter used to construct Fig. 1 and is shown to facilitate the comparison. The subtropical gyre of the North Pacific Ocean is readily apparent in the altimetric result. In the South Pacific, no such simple gyre has emerged—rather, we see something resembling a zonal flow field. Although the South Pacific is poorly covered by conventional hydrography, we note the

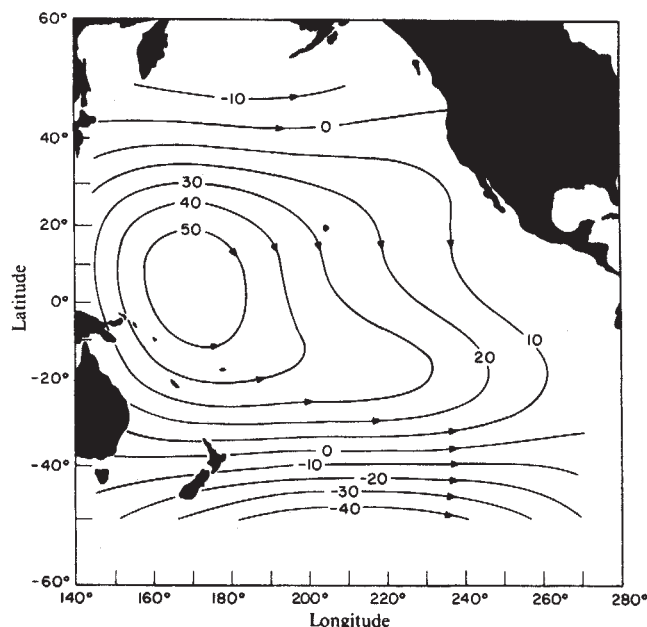


Fig. 3 As Fig. 2, the dynamic topography of the Pacific Ocean determined from shipboard hydrography, but run through the same low-pass wavenumber filter used to construct Fig. 1.

resemblance between Figs 1 and 3 in this region. All the evidence suggests there is probably no simple gyre in the South Pacific.

Both Figs 1 and 3 have geostrophic flow lines intersecting the coasts and crossing the Equator. The high wavenumber energy removed in the low-pass filtering operation is necessary to satisfy appropriate boundary conditions.

Two features in Fig. 1 are probably artefacts of remaining geoid error. The low in the north-east Pacific is too deep, causing the North Pacific gyre to bend too sharply to the south. In the south-east Pacific, the existence of the local high depicted there is doubtful. There is no quantitative proof of the impossibility of either of these features. The Wyrki chart, and others like it, should be interpreted with caution, for it is necessarily an inhomogeneous, aliased, representation obtained by combining the 70-yr historical record of shipborne hydrography.

If pictures such as Fig. 1 can be made from only 3 months of data from a no-longer state-of-the-art altimetric mission, and with an estimate of N that can be greatly improved with subsequent data, then we can see the possibility for studying the ocean as a complete entity for the first time with future optimal missions². Thus, we could determine such parameters as the multi-year means and the seasonal movement of water. An improved Fig. 1 could be used to correct Fig. 2 should we wish, or to constrain a general circulation model of the ocean. The use of measurements of ζ for discussions of the full three-dimensional oceanic circulation has been described elsewhere^{4,11}.

An extension of Fig. 1 to the world ocean is now being undertaken.

We thank R. Rapp for making available the Pacific result of his crossing-arc analysis. This work was supported in part by the US NASA under grant NAG6-9 and by the US NSF under grant OCE80-18514.

Received 20 September; accepted 25 November 1982.

1. Born, G. H., Dunne, J. A. & Lame, D. B. *Science* **204**, 1405–1406 (1979).
2. Cheney, R. E. & Marsh, J. G. *J. geophys. Res.* **86**, 473–483 (1981).
3. Bernstein, R. L., Born, G. H. & Whritner, R. H. *J. geophys. Res.* **87**, 3261–3268 (1982).
4. Wunsch, C. & Gaposchkin, E. M. *Rev. Geophys. Space Phys.* **18**, 725–745 (1980).
5. Topex Science Working Group, *Satellite Altimetric Measurements of the Ocean*, Jet Propulsion Laboratory, No. 400-111 3/8 (1981).
6. Rapp, R. H. *J. geophys. Res.* **84**, 3784–3792 (1979).
7. Marsh, J. G., Cheney, R. E., Martin, T. V. & McCarthy, J. J. *EOS* **63**, 178–179 (1982).
8. Lerch, F. J., Klosko, S. M., Laubacher, R. E. & Wagner, C. A. *J. geophys. Res.* **84**, 3897–3916 (1979).
9. Tai, C.-K. *J. geophys. Res.* (submitted).
10. Wyrki, K. *J. Phys. Oceanogr.* **5**, 450–459 (1975).
11. Munk, W. & Wunsch, C. *Phil. Trans. R. Soc. A* **307**, 439–464 (1982).

Direct determination of aluminium in natural waters by laser stepwise photoionization

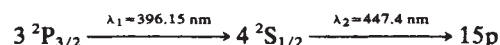
G. I. Bekov, A. S. Yegorov*, V. S. Letokhov & V. N. Radayev

Institute of Spectroscopy, USSR Academy of Sciences, 142092, Moscow Region, Troitzk, USSR
*Shirshov Institute of Oceanology, USSR Academy of Sciences, Moscow, USSR

The development of new, highly sensitive, methods of analysis of microelements and aluminium, particularly in natural waters, is important because: (1) aluminium, one of the most important elements in the lithosphere, is used in lithology and geochemistry of ocean as an indicator of terrigenous runoff; and (2) there are data^{1,2} relating aluminium dissolved in seawater to certain biological processes. However, the sensitivity of conventional analytical methods in most cases is insufficient for a direct determination of aluminium dissolved in sea and

ocean waters^{3,4}. Modern laser methods of single atom detection (see refs 5, 6) seem promising for this purpose due to their high detection efficiency; of particular interest is the method of laser stepwise photoionization of atoms^{7,8}. We describe here a new highly sensitive analytical method based on this technique which involves thermal atomization of the sample in vacuum and detection of the element atoms through their stepwise excitation to a Rydberg state and following ionization by electric field pulse.

Aluminium was determined in natural waters as follows. A water sample (40 μ l volume) was introduced into a tantalum crucible and evaporated at 90 °C to obtain a dry residuum. The crucible and contents were placed in a vacuum chamber, evacuated to a pressure of 10^{-6} torr and heated stepwise to 1,750 °C. The evaporating atoms and molecules of the sample were formed into an atomic-molecular beam using diaphragms. The beam was irradiated between two electrodes by two tunable dye lasers pumped simultaneously with a nitrogen laser. The wavelengths of the dye lasers were tuned for two-step excitation of aluminium atoms to a Rydberg state by the following process



After laser excitation of the atoms to a Rydberg state, 20 ns later an electric pulse that ionized all the Rydberg atoms was fed to the electrodes. The resulting ions were pushed by the same electric field pulse through a hole in one of the electrodes, then they were detected by an electron multiplier. The ion signal from the electron multiplier was passed to a boxcar averager (PAR model 162) and was registered by a recorder. In this experiment the total efficiency of ionization of Al atoms from the $3^2P_{3/2}$ state was determined by the efficiency of Rydberg state excitation and found to be 10%. Details of the experimental setup and the principle of its operation are given elsewhere^{6,9}.

Figure 1a illustrates the typical dependence of the ion signal on sample evaporation time at a temperature of 1,750 °C for 40 μ l of seawater. During 0–5 min the crucible was heated stepwise to 1,500 °C. In this temperature region the ion signal of Al was almost absent. Increasing the crucible temperature to 1,750 °C gave rise to an intense ion signal (the rate of crucible temperature build-up was not higher than 100 °C s⁻¹). To discriminate the selective ion signal of aluminium from the total ion signal the wavelength of the first-step laser was periodically tuned off by 3–5 cm⁻¹ from the aluminium transition $3^2P_{3/2} \xrightarrow{\lambda_1} 4^2S_{1/2}$. The dips in Fig. 1a correspond to the detuning of λ_1 from resonance. This detuning causes the ionization efficiency of aluminium atoms to drop by more than two orders of magnitude. Such a technique allowed the nonselective ion background of the seawater matrix to be discriminated from the ion signal conditioned only by aluminium. The total yield

Table 1 Content of dissolved aluminium in the zone of the mouth of the Kura river (the Caspian Sea)

No. of station	Sampling horizon (m)	Salinity (‰)	Al concentration* (μ g l ⁻¹)
1	0	<1	190 ± 20
	2	<1	215 ± 22
2	0	7.7	780 ± 80
	3	11.6	36 ± 4
3	0	10.1	230 ± 25
	10	12.2	260 ± 28
4	0	10.4	460 ± 50
	12	12.1	380 ± 40
5	0	12.4	230 ± 25
	22	12.3	135 ± 15
6	0	12.4	95 ± 10
	30	12.4	135 ± 15

* To a confidence level of 0.95.

of aluminium in this case is defined by the area lying between the curve of the total ion signal and the dashed line. The nonselective background is probably conditioned by thermal ionization of the easily ionized components of seawater and laser photoionization of the compounds constituting the seawater matrix.

The construction of an analytical calibration characteristic is a generally accepted technique in quantitative measurements. The calibration solutions in this case must be prepared on the basis of the matrix being studied. First we studied the yield of aluminium from known solutions, prepared on the basis of aluminium chloride and deionized water. The typical time dependence of the ion signal for the calibrating solution of AlCl_3 with the Al content being $100 \mu\text{g l}^{-1}$ is shown in Fig. 1b ($t_{\text{at}} = 1,750^\circ\text{C}$). In the case of AlCl_3 solution the ion signal is conditioned by aluminium only.

The additions method has shown that the matrix composition of seawater does not distort the results. We prepared a composite solution consisting of 9 parts of seawater and 1 part of calibrating solution of AlCl_3 with the Al content being $1,000 \mu\text{g l}^{-1}$. This corresponds to the additional concentration $100 \mu\text{g l}^{-1}$ of Al in seawater. The dependence of the ion signal on evaporation time for such a mixture in the same conditions

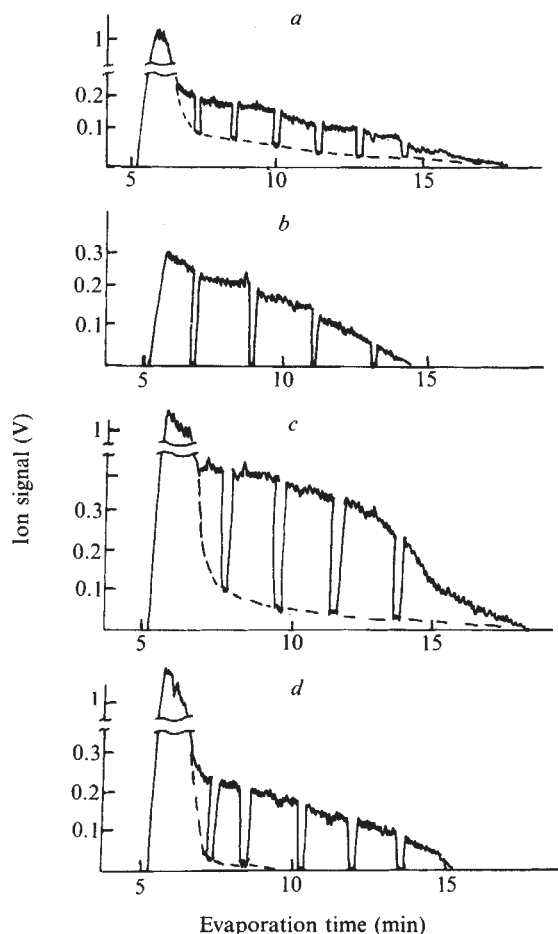


Fig. 1 Dependence of the ion signal on the time of sample atomization at $1,750^\circ\text{C}$: *a*, $40 \mu\text{l}$ of seawater. The dips are caused by reduction of detection efficiency of Al atoms by more than two orders of magnitude because the laser radiation frequency is detuned from the absorption line of Al. The total yield of Al is defined by the area between the curve of total ion signal and the dashed line. *b*, $40 \mu\text{l}$ of calibration AlCl_3 solution in deionized water (the concentration of Al is $100 \mu\text{g l}^{-1}$). *c*, $40 \mu\text{l}$ of solution containing of 9 parts of seawater and 1 part of calibration AlCl_3 solution with an Al content of $1,000 \mu\text{g l}^{-1}$. The total yield of Al is equal to the sum of the yields from *a* and *b*. *d*, $40 \mu\text{l}$ of aqueous solution with the concentrations of NaCl of $2 \times 10^7 \mu\text{g l}^{-1}$ (such concentration of NaCl is typical of seawater) and Al of $100 \mu\text{g l}^{-1}$.

The yield of Al is equal to that in case *b*.

is presented in Fig. 1c. The total signal of aluminium determined by the 'selective' area under the signal curve equals the selective signal of aluminium from seawater (Fig. 1a) plus the signal of aluminium from the solution with the concentration of Al being $100 \mu\text{g l}^{-1}$ (Fig. 1b). The study of the aqueous solution (Fig. 1d) containing $2 \times 10^7 \mu\text{g l}^{-1}$ of NaCl (such a concentration of NaCl is typical of seawater) and aluminium with its concentration of $100 \mu\text{g l}^{-1}$ has been attempted to analyse the nature of nonselective ion signal in seawater (Fig. 1a). The ion signal in Fig. 1d also consists of selective and nonselective parts. The area of the selective signal shows that the yield of Al here is similar to that of the solution of AlCl_3 (Fig. 1b), that is, NaCl has no effect on the yield of Al. The nonselective ion signal conditioned by ionization of NaCl (Fig. 1d) differs from the nonselective ion signal for seawater by its time behaviour (Fig. 1a, c).

The analytical calibration characteristic constructed on the basis of AlCl_3 solutions in deionized water is presented in Fig. 2a. It turned out to be linear in the concentration range $1-10^3 \mu\text{g l}^{-1}$. Above $10^3 \mu\text{g l}^{-1}$ the calibration characteristic became nonlinear because of saturation of the recording apparatus. The lower limit of the calibration characteristic was determined by the aluminium content in deionized water and

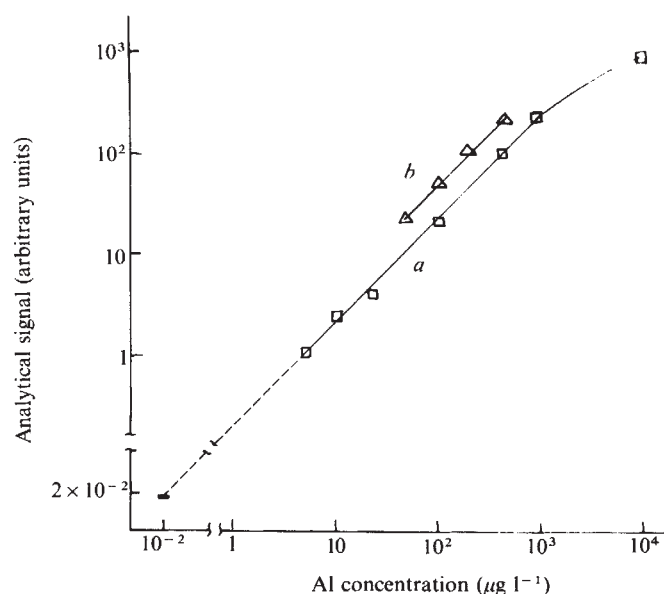


Fig. 2 Analytical calibration characteristics for Al on the basis of AlCl_3 solution in: *a*, deionized water; *b*, seawater. The dashed line denotes the calibration characteristic extrapolation to the noise level for finding the detection limit of Al that came to $0.01 \mu\text{g l}^{-1}$.

Table 2 Content of dissolved aluminium in the pore waters of the Indian Ocean

Station coordinates				Al concentration*
Lat.	Long.	Ocean depth (m)	Sediment type	($\mu\text{g l}^{-1}$)
8°45' S	102°19' E	5,300	Pelagic slime	17 ± 2
2°63' N	80°10' E	4,300	Clay slime	14 ± 2
0°04' N	80°02' E	4,600	Clay slime	30 ± 3
10°10' S	67°31' E	4,060	Carbonaceous slime	34 ± 4
9°08' S	67°16' E	6,015	Siliceous slime	29 ± 3
26°38' S	67°31' E	4,780	Metalliferous sediment	220 ± 25
23°54' S	73°56' E	3,780	Metalliferous sediment	54 ± 6
23°00' S	75°21' E	4,320	Metalliferous sediment	100 ± 11

* To a confidence level of 0.95.

came to $5 \mu\text{g l}^{-1}$. The detection limit of aluminium in these experiments determined by extrapolating the calibration characteristic to the background level was $0.01 \mu\text{g l}^{-1}$. Figure 2b shows the calibration characteristic constructed on the basis of seawater with AlCl_3 additions. The characteristics have proved to be parallel, which indicates the possibility of using the 'pure' calibration characteristic in quantitative measurements of aluminium in the seawater. The nonselective ion background affects only the value of minimum detectable concentration of aluminium in seawater (which was $1 \mu\text{g l}^{-1}$ in our experiments) as it is difficult to discriminate a small selective signal against the background of a strong nonselective one. The relative standard deviation determined from 20 parallel measurements of seawater samples came to 5.6%. The largest contribution to this value is from the variable amounts of sample introduced in the crucible.

The method described above has been used to measure dissolved aluminium in the water samples taken in the mouth of the Kura River (the Caspian Sea). The samples were taken with a 10-l bathometer and filtered through $0.45\text{-}\mu\text{m}$ membrane filters. The results of the measurements are given in Table 1.

Received 11 June; accepted 23 November 1982.

1. Mackenzie, F. T., Wollast, R. & Stoffyn, M. *Science* **199**, 4329 (1978).
2. Caschetto, S. & Wollast, R. *Geochim. cosmochim. Acta* **43**, 3 (1979).
3. Dickinson Burrows, W. *CRC Critical Reviews in Environmental Control*, 167-216 (1977).
4. Slavin, W. *Atom. Spectrosc.* **1**, 66 (1980).
5. Letokhov, V. S. in *Chemical and Biochemical Applications of Lasers* Vol. 5 (ed. Moore, C. B.) 1-38 (Academic, New York, 1980).

The results obtained for river water approximate the average value for the rivers throughout the world³.

Table 2 presents the results of dissolved aluminium concentration measurements in the pore water samples of the Indian Ocean. The pore water was squeezed out of some sediment monoliths (10-15 cm below the bottom) picked with a bottom scoop and filtered through $0.45 \mu\text{m}$ membrane filters. (The samples of pore waters were taken during the 1980 run of the research ship *Dmitry Mendeleev*.) The results obtained agree well with those in ref. 2.

The dissolved aluminium in the samples of surface waters of the Mediterranean has also been measured. The value obtained ($6.5 \pm 0.7 \mu\text{g l}^{-1}$) is close to the average concentration of aluminium in the Mediterranean Sea¹⁰.

Our experiments show that laser stepwise atomic photoionization in combination with thermal atomization of matter in vacuum is a promising method for determining traces of elements in natural waters.

We thank Professor A. P. Lisitsyn for his support, Professor O. V. Shishkina for useful discussions and R. Mamed-Zade for providing water samples from the river-sea zone.

6. Balykin, V. I., Bekov, G. I., Letokhov, V. S. & Mishin, V. I. *Usp. Fiz. Nauk* **132**, 293-344 (1980); *Soviet Phys. Usp.* **23**, 651-678 (1980).
7. Letokhov, V. S. USSR Patent No. 784679 (Appl. 30.03.1970; Publ. 1.08.1980).
8. Ambartsumyan, R. V., Kalinin, V. P. & Letokhov, V. S. *Pis'ma ZhETF* **13**, 305 (1971); *JETP Lett.* **13**, 217 (1971).
9. Bekov, G. I., Vidolova-Angelova, E. P., Letokhov, V. S. & Mishin, V. I. in *Springer Series in Optical Sciences*, Vol. 21 (eds. Walther, H. & Rothe, K. W.), 283-294 (Springer, Berlin, 1979).
10. Caschetto, S. & Wollast, R. *Mar. Chem.* **7**, 2 (1979).

Mudrocks examined by backscattered electron microscopy

K. Pye

Department of Earth Sciences, University of Cambridge,
Downing Street, Cambridge CB2 3EQ, UK

D. Krinsley

Department of Geology, Arizona State University, Tempe,
Arizona 85287, USA

The small size of many of the particles in mudrocks makes it almost impossible to image and identify them individually and *in situ* using conventional light microscopy. Partly for this reason, the petrography and genesis of mudrocks is poorly understood compared with sandstones and limestones. This deficiency presents a serious problem in sedimentary geology, as mudrocks comprise almost two-thirds of the stratigraphical column^{1,2}. We introduce here a new method of studying mudrocks, using backscattered electrons (BSE) in scanning electron microscopy (SEM). When thin sections or polished rock chip surfaces are examined with BSE, both atomic number (*Z*) contrast and topographic contrast can be obtained, providing much more detailed compositional and textural information than was previously possible. Used in conjunction with X-ray microanalysis, the new technique promises to revitalize mudrock petrology.

Secondary electrons in SEM have previously been used to study geological specimens, but for various reasons, this technique has not been widely applied or found to be very useful in the analysis of mudrocks. First, fracture surfaces of broken specimens have conventionally been studied rather than thin sections or polished surfaces. The general belief has been that because secondary electrons image topography, rough, broken surfaces should provide maximum information. However, in practice when fracture surfaces of fine-grained materials such as mudrocks are examined with secondary electrons, individual mineral grains cannot normally be distinguished, and the textural information obtained is of limited value. Second, although

it has been known for several years that backscattered electrons can produce *Z* contrast³⁻⁵, the Everhart-Thornley electron detectors fitted as standard on most SEMs were capable of producing only poor resolution BSE images (by introducing a negative grid voltage on the Faraday cage). Recently, several detectors designed specifically to collect backscattered electrons have been made commercially available. These are capable of producing images having comparable or superior resolution to secondary electron images. With the more sophisticated systems, it is possible to mix the BSE and secondary electron signals to obtain the optimum image for a particular material. We illustrate here how these detectors can be used to provide very detailed information about the form and composition of individual grains in mudrock thin sections. This information can be used, for example, in studies of the source, mode of deposition, diagenesis and tectonic deformational history of mudrocks.

Uncovered thin sections of an Upper Ordovician mudstone from Gwynedd, Central Wales, and an organic-rich Devonian mudstone from Pennsylvania, USA, were prepared in the normal way and then polished as finely as possible. Polishing is desirable to reduce topographic contrast and optimize the *Z* contrast observed when the specimen is examined with backscattered electrons⁵. After mounting on SEM stubs, the sections were coated with $\sim 100 \text{ \AA}$ of carbon to eliminate charging. A thin trail of silver dag was run from the top surface of the specimen to the stub to improve electrical conductivity. This is essential, as poor conductivity severely limits resolution.

The sections were examined with an ISI SS40 SEM equipped both with a Taylor directional-type BSE detector and a solid-state annular BSE detector. Figure 1 illustrates the *Z* contrast obtained from the Welsh Ordovician mudstone using the solid-state detector. As this detector was mounted directly above the untitled specimen, and all four quadrants of the detector were switched on, no topographic contrast is seen in Fig. 1. Energy dispersive X-ray microanalysis confirmed that the differences in grey level correspond to variations in mineralogical composition. The differences in grey level arise because the backscattering coefficient (η) increases as a function of the average atomic number (*Z*) of the target^{5,6}. Minerals containing large amounts of relatively heavy elements, such as pyrite, rutile, ilmenite or zircon, appear light, while aluminosilicates such as illite or



Fig. 1 BSE micrograph of a polished thin section of an Upper Ordovician mudstone, taken using a solid-state detector. Minerals identified by X-ray microanalysis include: (1) chlorite; (2) illite; (3) pyrite; (4) quartz. This micrograph shows only atomic number (Z) contrast. Scale bar, 30 μm .

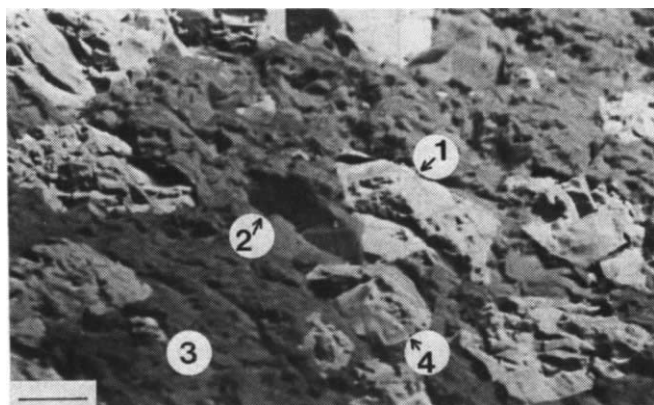


Fig. 2 BSE micrograph of a polished thin section of Devonian mudstone, taken using a Taylor scintillator-photomultiplier detector and specimen tilt of 40° . Mineral phases present are: (1) pyrite; (2) K-feldspar; (3) illite; (4) quartz. This micrograph shows both Z-contrast and topography. Scale bar, 20 μm .

feldspars appear relatively dark. The large feature on the left-hand side of Fig. 1 is a chlorite-mica stack⁷, consisting largely of chlorite (light) with thin, irregular interlayers of illite (dark bands). Electron microprobe analysis has indicated the composition of the chlorite to approximate that of ripidolite, while the mica appears to be a potassium-poor, hydrous form of illite⁸. Smaller chlorite-mica stacks are scattered about the micrograph. The matrix consists largely of illite plates, 2–15 μm long and 1–2 μm thick, which show a crude preferred orientation parallel to cleavage. Quartz and feldspar grains are also present. Although they show similar grey levels to illite, they can be distinguished on the basis of their form, supported by X-ray microanalysis. Individual pyrite cubes and framboids show up white due to their relatively high average atomic number. The black areas are voids which may represent natural porosity. Many of the grains, particularly the quartz and feldspar, show sharp, straight edges, sutured inter-grain contacts and re-entrant features, suggesting that the particles have experienced pronounced post-depositional diagenetic modification. Grain realignment appears relatively weak, however, and there is little evidence of severe tectonic deformation. In a systematic analysis of strain patterns, the size, shape, orientation and distribution of different minerals could easily be mapped, digitized and quantified. Similarly, the distribution and form of void spaces could be examined by image analysis. X-ray maps of particular elements could also be prepared using the X-ray microanalysis unit in the scan mode.

Although Z contrast is optimized when topographic contrast is reduced to a minimum, it is sometimes advantageous when studying the fine detail of mudrock fabrics to produce micrographs which combine Z contrast with topographic information. Topography can be obtained in various ways, including shutting off two of the quadrants of the solid-state detector, or by tilting the specimen relative to a directional BSE detector such as the Taylor detector. Figure 2 is a micrograph of the Devonian mudstone taken using the Taylor detector and with the specimen tilted at $\sim 40^\circ$. The topography is evidently related to differences in hardness which are exploited during polishing of the specimen. Quartz, feldspar and pyrite appear to stand out above the illite-rich matrix. The combination of topographic and atomic number contrast can help in the interpretation of mudrock fabric when there is a significant difference in the average atomic number of the individual mineral constituents. Where this is not the case, however, it may be more informative to compare separate Z-contrast and topographic images of the same area. In most investigations, Z-contrast-only images and Z-contrast plus topography images should be complementary.

We believe that the application of BSE in SEM studies of mudrock thin sections and polished rock chip surfaces will provide the shale petrographer with a tool of equivalent value to the polarizing microscope in sandstone and carbonate petrology. Used in conjunction with X-ray microanalysis, the distribution of specific minerals in mudrocks can be mapped throughout a thin section and related to cleavage development or diagenetic stage. Intrastratal solution, overgrowths, grain-to-grain contacts, relative hardness and relative ages of diagenetic growth stages can be studied. In some cases it may also be possible to make inferences about provenance and mode of deposition on the basis of grain size, shape, composition and texture. Thus the technique offers the basis for a 'new shale petrography' and should ultimately lead to much better understanding of the internal structure and genesis of mudstones.

K.P. acknowledges a NERC Research Fellowship. D.K. wishes to acknowledge NASA-Ames research grant NAG 2-174 and the Planetary Geology Program, Office of Space Science, Research grant NAGW24. We thank Brian Krinsley for assistance. Earth sciences contribution 331.

Received 25 October; accepted 25 November 1982.

1. Potter, P. E., Maynard, J. B. & Pryor, W. A. *Sedimentology of Shale* (Springer, New York, 1980).
2. Blatt, H. *Sedimentary Petrology* (Freeman, San Francisco, 1982).
3. Robinson, V. N. E. *Scanning* 3, 15–26 (1980).
4. Ball, M. D. & McCartney, D. G. *J. Microsc.* 124, 57–68 (1981).
5. Hall, M. G. & Lloyd, G. E. *Am. Miner.* 66, 362–368 (1981).
6. Goldstein, J. I. et al. *Scanning Electron Microscopy and X-ray Microanalysis* (Plenum, New York, 1981).
7. Craig, J., Fitches, W. R. & Maltman, A. J. *Geol. Mag.* 119, 243–256 (1981).
8. Krinsley, D., Pye, K. & Kearsley, A. T. *Geol. Mag.* (in the press).

Glacial erosion in an ice-divide zone

R. Cornish

Department of Economics, Heriot-Watt University,
Edinburgh EH1 2HT, UK

It is widely accepted that glacial erosion, and therefore deposition, are insignificant at the centres of ice-sheets^{1–4}, although a few authors have argued for extensive glacial erosion in those areas^{5,6}. Here I argue from the detailed study of erratics trains from three distinctive bedrock sources, from striae and from ice-moulded landforms that the two views need not be mutually exclusive. The presence of large quantities of lodgement till in the vicinity of a former ice-divide in south-west Scotland is explained by debris emplacement during ice-sheet build-up, when conditions for erosion obtained. The till was protected from subsequent removal by the superimposition of an ice-divide across the area.

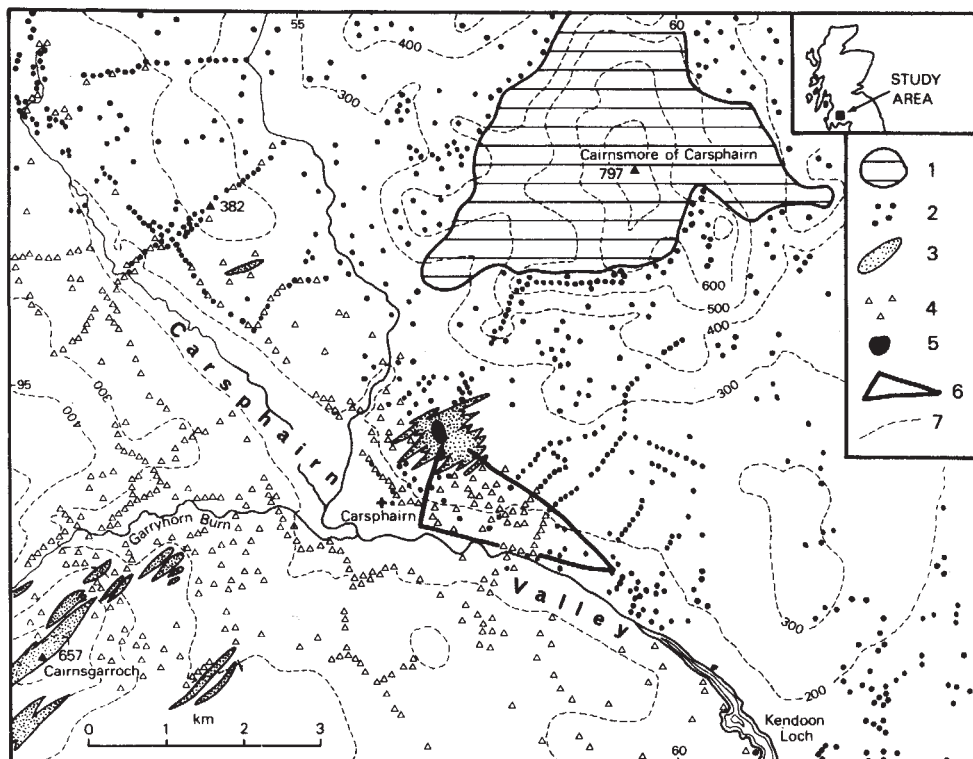


Fig. 1 Distribution of erratics in the Carsphairn Valley: 1, Cairnsmore granitic intrusion; 2, granite erratics; 3, Ordovician conglomerate outcrops; 4, conglomerate erratics; 5, Knockgray microgranite intrusion; 6, envelope defining distribution pattern of microgranite erratics; 7, contours (100-m interval).

The Loch Doon basin in the western Southern Uplands of Scotland was a major centre of ice dispersal during ice-sheet glaciations⁷. The distribution of granite erratics from the basin has been mapped and the existence of a former south-west-north-east-trending ice-divide has been demonstrated^{8,9}. The Carsphairn Valley lies to the east of the Loch Doon basin and comprises a through-valley trending north-west-south-east, which is flanked on both sides by uplands reaching 700 m in altitude (Fig. 1). The central part of the Carsphairn Valley is characterized by extensive deposits of till (Fig. 2). The till comprises both elongated ridges (which occur in the Carsphairn Valley and trend perpendicular to the valley axis, but are drumlinized in a direction parallel to the valley axis) and amorphous valley-fill deposits that choke the side valleys (Garryhorn Burn for example, Fig. 1). The former merge into drumlins to the north-west and south-east. The ridges reach a maximum height of 20 m, while the valley-fill is at least 18 m thick. The till comprising both types of landform is overconsolidated, shows fissile structure and no sign of interbedding with sand and gravel. All these characteristics indicate a subglacial origin for the till and infer emplacement by lodgement processes.

Former directions of ice flow in the Carsphairn Valley were determined by mapping the distribution of erratics from three bedrock sources, namely Cairnsmore granite, Knockgray microgranite and Ordovician conglomerates. Since most of the mapped erratics occur on the ground surface, emplacement by the last ice-sheet to cover the area (late Devensian) is inferred. Supplementary information on ice movement direction was provided by mapping striae and ice-moulded landforms.

Striae and ice-moulded landforms in the northern part of the Carsphairn Valley indicate former ice flow from south-east to north-west, while those in the southern part show a north-west to south-east trend. This pattern of flow is confirmed by the distribution of conglomerate erratics (Fig. 1): these are also of limited distribution, the farthest-travelled conglomerate boulder being only 4.4 km from the nearest outcrop. The conglomerate erratics also reveal the former presence of an ice-divide zone in the Carsphairn area, because they occur both north-west and south-east of the outcrops. This conclusion is supported by the distribution of Cairnsmore granite erratics, which show a similar pattern (Fig. 1). The former ice-divide is therefore placed as crossing the Carsphairn Valley from Cairnsgarroch

summit through Craig of Knockgray to the Cairnsmore hills (Fig. 2). This is also the area where the lodgement till is most extensive and reaches its greatest thickness.

The microgranite erratics train is very short, the farthest-travelled boulder occurring only 2.8 km from the source. These erratics occur only on the south and south-east slopes of Craig of Knockgray (Fig. 1), but subtend the unusually large angle of 90°. This indicates that directions of ice flow over the outcrop varied greatly during emplacement of the erratics. The microgranite outcrop lies immediately north-west of Craig of Knockgray summit and is surrounded by conglomerate. To account for the absence of microgranite erratics from the northern slopes of the hill, the former ice-divide must be placed over the conglomerate outcrop to the north-west of the microgranite outcrop. This allows the ice-divide to be located to within ± 300 m in this section of the valley.

The microgranite erratics train is significant in that its location beneath the former ice-divide zone permitted only limited dispersal of boulders, as horizontal ice movement at an ice-divide is theoretically zero¹⁰ and the capacity for ice to erode eliminated¹¹. The limited distribution also indicates that there was no migration of the ice-divide. However, the large quantities of lodgement till in the Carsphairn area, across the former ice-divide zone, require explanation. As the till cannot have been deposited when the ice-divide existed, due to zero ice movement, emplacement of the debris occurred either before or after the ice-divide came into existence. Emplacement during the early stages of ice-sheet decay is rejected for two reasons. First, the unusual distribution pattern of microgranite erratics can only be explained if ice first flowed southwards, towards the axis of the Carsphairn Valley, this being followed by establishment of the ice-divide with limited movement towards the south-east. Second, the fact that mapped striae in the area display a clear pattern away from the ice-divide implies that after the ice-divide existed there was no period when ice flowed into the Carsphairn Valley from upland source areas on both sides.

Glaciers that existed during the Loch Lomond Stadial (approximately equivalent to the Younger Dryas) have been mapped in the valleys on each side of the Carsphairn Valley^{12,13}. These former glaciers are of very restricted extent, being confined to corries and trough-head locations. In some places

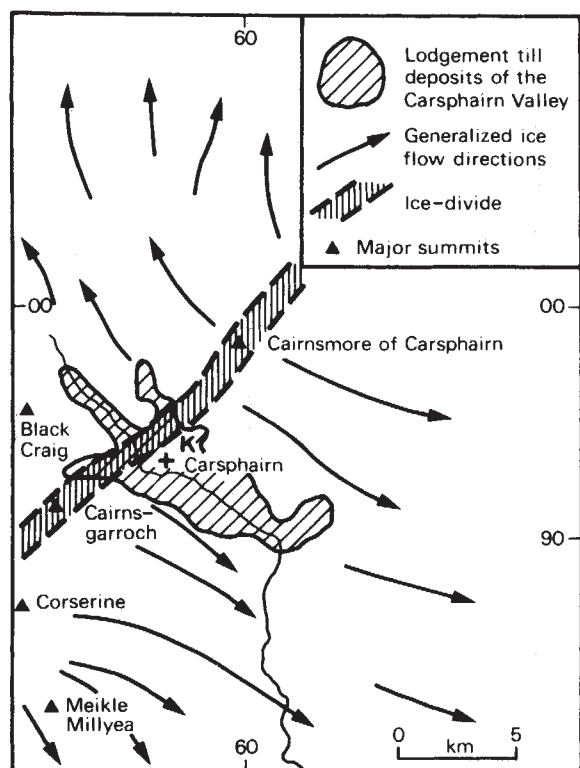


Fig. 2 Generalized ice flow directions, position of the former ice-divide and lodgement till deposits in the Carsphairn area. K, Craig of Knockgray summit.

(in the Garryhorn Burn Valley) the glacier limits show that the former glaciers ploughed into the margins of the late Devensian till deposits to produce end moraines, demonstrating that the stadial glaciers developed after the till was deposited. This is supported by coleopteran evidence from south-west Scotland¹⁴, which showed that at 13,000 yr BP the mean July temperature was $\sim 15^{\circ}\text{C}$ and that the climate was little different from the present. Such evidence suggests that, in this part of Scotland, climatic conditions during the Lateglacial Interstadial (approximately equivalent to the Allerød) were not conducive to glacier survival after ice-sheet deglaciation. These observations indicate a recrudescence of ice in sheltered locations during the severe climatic deterioration of the Loch Lomond Stadial. Thus the glaciers that formed during this period may be used to portray an early stage in ice-sheet growth, as they identify likely areas of initial ice build-up. Had the Loch Lomond Stadial lasted longer the glaciers described above, and perhaps others, would have expanded onto lower ground, ultimately converging and coalescing in the Carsphairn Valley. Such flow is demonstrated by the presence of Cairnsmore granite erratics in the Carsphairn area (south-west of the outcrop) and by conglomerate erratics emanating north-east from the Garryhorn Valley (Fig. 1). I suggest that glaciers moving from these opposing directions came into contact, causing flow to be impeded and debris build-up to occur in the central Carsphairn Valley. With subsequent expansion of the ice-sheet and development of the ice-divide, the debris and erratics escaped removal owing to very slow rates of ice movement.

It has been generally assumed that maximal glacial erosion is associated with maximal development of an ice-sheet, and that erosion occurred in areas away from the centre where the basal freezing-on mechanism was operative^{2,15,16}. However, the above discussion suggests that deposition of thick lodgement till and therefore, by implication, that significant glacial erosion occurred during the earlier stages of ice-sheet growth. These conclusions imply that high erosion rates may occur beneath ice-sheet source areas, not during ice-sheet maxima, but during ice-sheet growth when higher velocities and necessarily, tem-

perate basal ice conditions pertained. Extensive areas of the central Laurentian Shield, western Scandinavia and the Scottish Highlands display evidence of intense glacial erosion, yet much of these areas lay beneath former ice-divides^{7,17,18}. It is concluded that erosion of this nature could only have occurred during ice-sheet growth or during periods of limited glaciation.

The preservation of till at the centre of an ice-sheet has implications for such subglacially deposited landforms as Rogen moraines, which commonly occur in these areas^{19,20}. Emplacement of debris before the establishment of an ice-divide would make the till available for the formation of the above landforms. This would obviate the need for the complex mechanisms of debris emplacement that have been invoked previously²¹⁻²³.

I thank Drs J. B. Sissons, D. G. Sutherland and M. J. C. Walker for helpful comments.

Received 18 August; accepted 26 November 1982.

1. Flint, R. F. *Glacial and Pleistocene Geology* (Wiley, New York, 1971).
2. Gravenor, C. P. *Am. J. Sci.* **275**, 594-604 (1975).
3. Sugden, D. E. *Geology* **4**, 580-582 (1976).
4. Andrews, J. T. *Quat. Res.* **17**, 123-124 (1982).
5. White, W. A. *Bull. geol. Soc. Am.* **83**, 1037-1056 (1972).
6. Laine, E. P. *Quat. Res.* **14**, 188-198 (1980).
7. Sissons, J. B. *The Geomorphology of the British Isles: Scotland* (Methuen, London, 1976).
8. Charlesworth, J. K. *Trans. R. Soc. Edin.* **55**, 1-23 (1926).
9. Cornish, R. thesis, Univ. Edinburgh (1979).
10. Paterson, W. S. B. *The Physics of Glaciers* (Pergamon, Oxford, 1981).
11. Boulton, G. S. in *Glacial geomorphology* (ed. Coates, D. R.) 41-87 (New York State University Press, Binghamton, 1974).
12. Cornish, R. *Proc. geol. Ass.* **92**, 105-114 (1981).
13. Holden, W. thesis, Univ. Glasgow (1977).
14. Bishop, W. W. & Coope, G. R. in *Studies in the Scottish Lateglacial Environment* (eds Gray, J. M. & Lowe, J. J.) 61-88 (Pergamon, Oxford, 1977).
15. Sugden, D. E. *J. Glaciol.* **21**, 367-391 (1978).
16. Nobles, L. H. & Weertman, J. in *Till: a Symposium* (ed. Goldthwait, R. P.) 117-126 (Ohio State University Press, Columbus, 1971).
17. Shilts, W. W. *Nature* **286**, 213-218 (1980).
18. Vorren, T. O. *Boreas* **6**, 247-257 (1977).
19. Lee, H. A. *Geol. Surv. Bull. Can.* **51**, 1-42 (1959).
20. Lundqvist, J. *Sver. geol. Unders.* **C648**, 1-32 (1969).
21. Sugden, D. E. & John, B. S. *Glaciers and Landscape* (Arnold, London, 1976).
22. Aario, R. *Geojournal* **1**, 65-72 (1977).
23. Shaw, J. *Boreas* **8**, 409-426 (1979).

Barrier dynamics and landward migration with Holocene sea-level rise

Stephen P. Leatherman

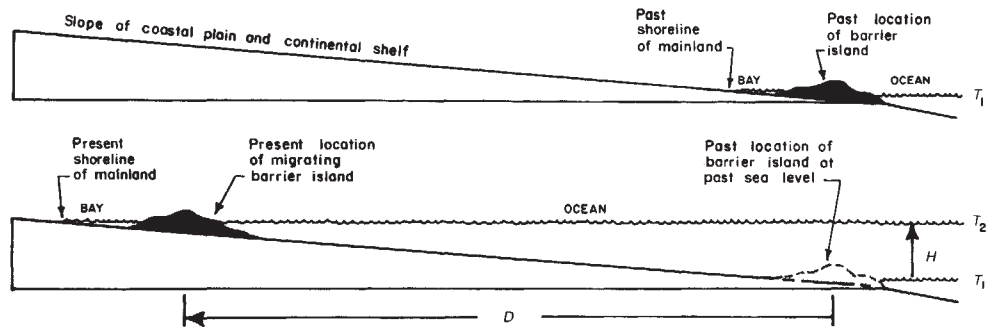
Department of Geography, University of Maryland, College Park, Maryland 20742, USA

Barrier beaches (islands and spits) have become the subject of mounting interest^{1,2} due to human development on these landforms. It is commonly thought that barrier beaches continuously migrate landward in response to sea-level rise, often by overwash processes, maintain mass through time, and that these processes are occurring at a rate commensurate with human (100 yr) time frames. While some barriers are naturally migrating quite rapidly^{3,4}, other barriers decidedly are not. Here I present data which illustrate that many US Atlantic Coast barriers are not evolving over the short term in accordance with generally accepted hypotheses and argue that the presumed norm may actually be the exception.

Geomorphical and historical evidence shows that most barriers are presently experiencing seaside erosion. In fact, shore erosion is a prevailing worldwide phenomenon along sandy coasts, except where ample sand supplies are readily available⁵. Initiation of the current transgressive phase is not precisely known, but most scientists agree that it began within the past 200 yr (refs 6,7). The forcing function for landward shoreline retreat is sea-level rise^{8,9}.

Prevailing theory holds that barrier response to sea-level rise is landward and upward migration through time^{10,11} (Fig. 1). Without sand movement from the ocean beach to the bayshore, the barrier feature would be forced to drown in place¹². Proponents of the overwash hypothesis¹³ argue that overwash is the principal process of landward migration, maintaining barrier integrity and relative elevation with respect to a rising sea.

Fig. 1 Within a geological time frame, barriers must migrate landward or be drowned in-place with sea-level rise. Large-scale barrier migration is associated with small increases in sea level due to the low, gradual slope of the coastal plain. This article defines the short-term dynamics of this long-term, time averaged phenomenon. H , rise in sea level; D , horizontal migration of barrier island where D is 100–1,000 times H ; T_1 , past sea level; T_2 , present sea level.



More recently, the predominant role of inlets in landward sediment transfers has been recognized^{14,15}.

There is geomorphic logic for this theory of barrier dynamics, but the precise mechanisms and rates for barrier rollover have not been properly assessed. Much speculation exists on this point at present; for instance, some scientists¹⁶ envision that barrier translocation can occur extremely rapidly (within 50–100 yr). Implicit in their basic argument is that barriers move as a virtual unit with conservation of mass.

Quantitative studies (ref. 17 and work in preparation) have shown that many barriers are actually eroding on both sides and experiencing bayside submergence. Fire Island (Fig. 2), a barrier island along the south shore of Long Island, New York, has been studied in detail (refs 18, 19 and work in preparation). Geomorphic information from >100 shallow vibracores indicate that the western half of the island has been relatively stable in its centroid position during at least the past millenia. Radiocarbon dates on *in situ* salt marsh peat deposits yield ages greater than 1,000 yr at relatively shallow depths (1.5–2.5 m below ground surface) near or at the barrier bayshore. As the depth of the enclosed bay (Great South Bay) ranges from 1.2 to 3.0 m, it appears that essentially no landward migration has taken place in this section of Fire Island during this time period¹⁹.

Stratigraphical correlation of dated peat layers at depth along bay to ocean transects also show that the salt marsh was displaced seaward during the past 500 yr. If the barrier were experiencing little or no landward migration, then sea-level rise would cause submergence of the backbarrier flats, which would be reflected in the stratigraphical record as seaward encroach-

ment of salt marshes towards the barrier nucleus. Other geomorphic evidence, including truncated inlet ridges and tree stumps *in situ* on the barrier bayside, provide conclusive evidence of bayside erosion and barrier submergence.

Landward barrier migration is a time-averaged phenomenon due to sporadic and site-specific storm-generated events over the long term. Radiocarbon dating of peat deposits encountered at depth show a discontinuous record of deposition. For instance, at one site along Fire Island (Fig. 2, point A) only 0.8 m of sediment were deposited between 1,350 and 350 yr BP, while 1.1 m of sediment accumulated during the past 350 yr. Other cores containing datable material illustrate similar chronologies and sequences. This barrier island is clearly not continuously migrating landward during human (<100 yr) time frames in response to sea-level rise.

The historical record confirms and provides detail on this geomorphic trend. Quantitative shoreline mapping shows that between the 1830s and 1979 bayside erosion equalled ocean-side erosion along some sections of Fire Island (Fig. 2). Therefore, the barrier is eroding from both sides, narrowing through time, rather than migrating landward. Previous studies at Assateague Island, Maryland¹⁵ showed that barriers must narrow to some critical width before initiation of migration. In the case of Assateague, emplacement of updrift jetties has sand-starved the northern end of this island, resulting in extremely high erosion rates exceeding 10 m yr⁻¹. This artificially-manipulated island continued to erode in place until the critical barrier width (200–250 m in this locality) was reached. At this point overwash is effective in transporting enough material over the island to compensate for shoreline losses so that a low, narrow

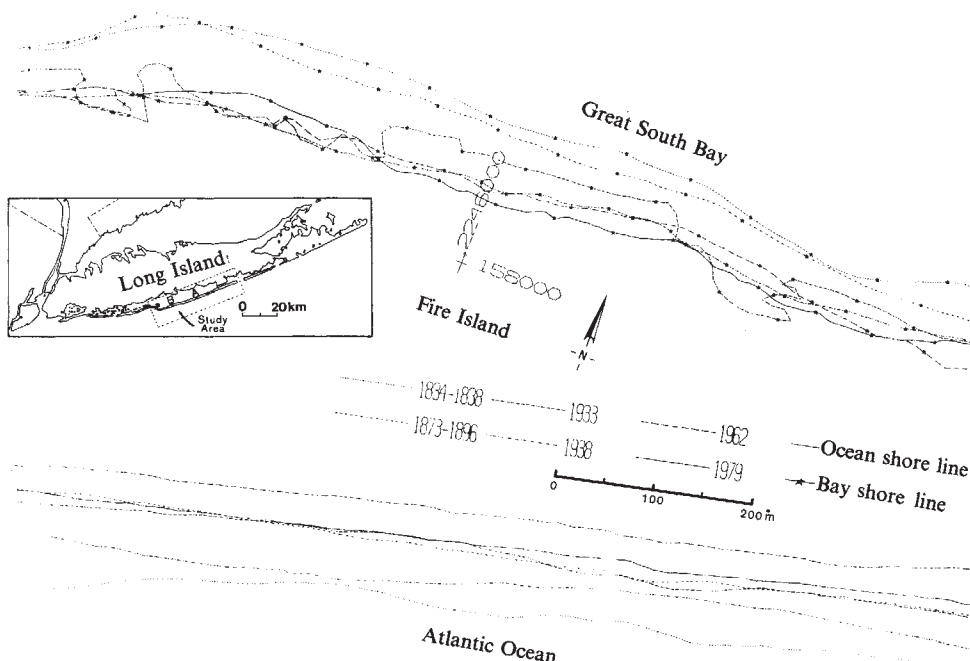


Fig. 2 A new, automated technique of mapping was utilized to generate this computer-plotted representation of the ocean and bay shorelines for the Sunken Forest area of Fire Island, New York (located at Point B on insert). The barrier centroid has remained essentially static during the past 150 yr as both shorelines exhibit an erosional trend.

barrier is being maintained with landward translocation. While ephemeral inlets and overwash to a lesser degree are effective in barrier rollover in these special conditions, the island section so affected has apparently not maintained mass as it has been drastically reduced in subaerial size with migration.

Quantitative studies along the Outer Banks of North Carolina (ref. 17 and C. Everts, work in preparation) confirm the Fire Island results and indicate the general applicability of these findings to a wide geographical area. During a 94-yr period (1852–1946), there was no substantial sand deposition along the bayside by overwash processes, and the barrier experienced both ocean and bayside erosion along much of its length. There was also a general decrease in total subaerial landmass such that the island is continuing to narrow through time and is apparently not maintaining volume. Due to the relative rise in sea level, Hatteras Island, North Carolina, is being submerged, producing favourable conditions for salt marsh growth towards the barrier nucleus; significant new salt marshes have developed on the backbarrier flats by seaward expansion. By comparison, overwash processes have been largely insignificant in providing new sandy substrate for salt marsh colonization. Note that this microtidal barrier system has been cited as a classic washover barrier^{13,16}, wherein overwash was believed to be chiefly respon-

sible for landward migration and salt marsh development during short time frames (<100 yr)²⁰. In practice, coastal investigators have often relied on qualitative data and observations of local events to formulate general models of barrier migration without recognition and full appreciation of transport mechanisms and time frames.

It is clear that barriers, such as Fire Island, New York and Hatteras Island, North Carolina have been essentially static for quite a long period of time (up to 1,000 yr), and migration is dominated by aperiodic inlet breaching¹⁵ occurring in a discontinuous step-like fashion. The evidence presented here has been derived from the well-studied US Atlantic Coast barriers, but these findings may have worldwide application in similar environmental settings.

It is not certain whether the present rise in sea level²¹ is a momentary upward blip on an essentially flat curve or the initiation of a major, long-term transgression. What is becoming increasingly evident is that many islands are eroding on both sides, not actually migrating as popularly believed. These barrier beaches will probably continue to 'slim down' for at least some time in the future; even if sea level stabilizes, the time lag in response variables will entail erosion and submergence for a period lasting from at least decades to a century.

Received 1 September; accepted 23 November 1982.

1. Leatherman, S. *Barrier Islands* (Academic, New York, 1979).
2. Leatherman, S. *Overwash Processes* (Hutchinson Ross, Pennsylvania, 1981).
3. Leatherman, S. & Zaremba, R. *U.S. Army Corps of Engineers Tech. Pap.* (in the press).
4. Penland, S. & Boyd, R. *Oceans* 1981, 209–219 (1981).
5. Bird, E. *Proc. 23rd Int. Geogr.* (1976).
6. Bakker, J. *Deutsch. Geogr. Verh.* 31, 233–237 (1957).
7. Tanner, W. *Geotimes* 27, 16–17 (1982).
8. Hicks, S. J. *geophys. Res.* 83, 1377–1379 (1978).
9. Emery, K. *Proc. natn. Acad. Sci. U.S.A.* 77, 6968 (1980).
10. Kraft, J. *Bull. geol. Soc. Am.* 82, 2131–2158 (1971).

11. Leatherman, S. *Barrier Island Handbook* (National Park Service, Boston 1979).
12. Rampino, M. & Sanders, J. *Sedimentology* 28, 37–47 (1981).
13. Dolan, R., Godfrey, P. & Odum, W. *Am. Scient.* 61, 152–162 (1973).
14. Armon, J. in *Barrier Islands* (ed. Leatherman, S. P.) 65–80 (Academic, New York, 1979).
15. Leatherman, S. *Geology* 7, 104–107 (1979).
16. Pilkey, O. & Myers, C. *Earth Sci.* 34, 20–23 (1981).
17. Jarrett, T. *U.S. Army Corps of Engng Rep.* (1981).
18. Leatherman, S. & Joneja, D. *National Park Service Tech. Rep. No. 47* (Boston, 1980).
19. Joneja, D., M. S. thesis, Univ. Massachusetts (1981).
20. Godfrey, P. *U.S. National Park Service Rep.* (1970).
21. Rampino, M. *Nature* 280, 132–134 (1979).

Prehistoric Maya wetland agriculture and the alluvial soils near San Antonio Rio Hondo, Belize

Paul R. Bloom*, Mary Pohl†, Cynthia Buttleman*, Frederick Wiseman‡, Alan Covich§, Charles Miksicek||, Joseph Ball¶ & Julie Stein#

* Department of Soil Science, University of Minnesota, St Paul, Minnesota 55108, USA

† Department of Anthropology, Florida State University, Tallahassee, Florida 32306, USA

‡ Department of Geography and Anthropology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

§ Department of Zoology, University of Oklahoma, Norman, Oklahoma 73019, USA

|| Department of General Biology, University of Arizona, Tucson, Arizona 85721, USA

¶ Department of Anthropology, San Diego State University, San Diego, California 92182, USA

Department of Anthropology and Quaternary Research Center, University of Washington, Seattle, Washington 98195, USA

The prehistoric Lowland Maya supported an elaborate civilization, which reached peak population densities in the tropical forests of Guatemala, Belize and southern Mexico during the Classic period between AD 300 and 900. It has been proposed that the Maya turned to labour-intensive construction of raised planting platforms in wetlands for year-round cultivation. Excavations at San Antonio Rio Hondo, northern Belize¹, provided the first archaeological data for this hypothesis. In our analysis of data from the San Antonio site, we conclude that the only indisputable evidence that prehistoric wetland cultivation occurred comes from a time before peak population of the Classic period, and that the technique required less labour than previously envisioned. Exploitation of swampy terrain initially involved dry-season cropping without extensive ditching and later ditching, probably for soil drainage.

The study site is located on Albion Island 2 m above sea level in the floodplain of the secondary channel of the Rio Hondo at the north-east margin of the village of San Antonio (18°10' N, 88°49' W). The soil parent material in the floodplain is calcareous alluvial clay. In the uplands of Albion Island, parent materials are well indurated limestone, poorly indurated limestone and chalk, which locally contain finely divided quartz and gypsum². Climate is tropical wet-dry according to the Koeppen classification. Annual rainfall is 1,500 mm.

In 1973 and 1974, Puleston's excavations (including 2B, 2J and 2W in Fig. 1) revealed platforms 12–20 m wide separated by narrow ditches 0.5–1.0 m wide. In 1980 we obtained soil,

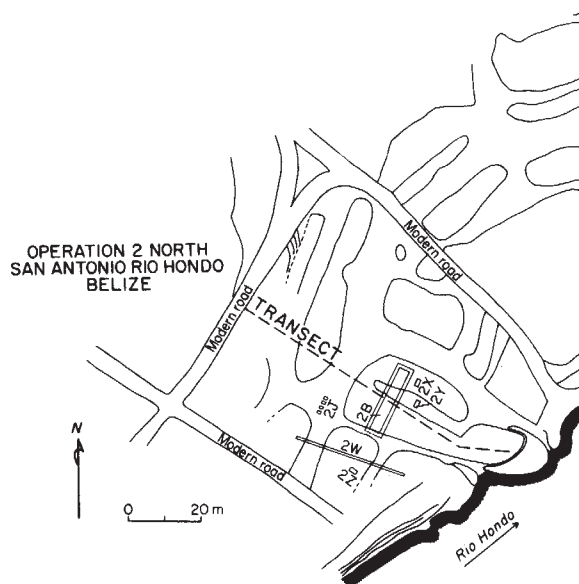


Fig. 1 Map of study area at San Antonio Rio Hondo, northern Belize. Pits 2B, 2J and 2W excavated by Puleston in 1973 and 1974. Pits 2X, 2Y, 2Z and transect by Bloom, Pohl and Buttleman in 1980.

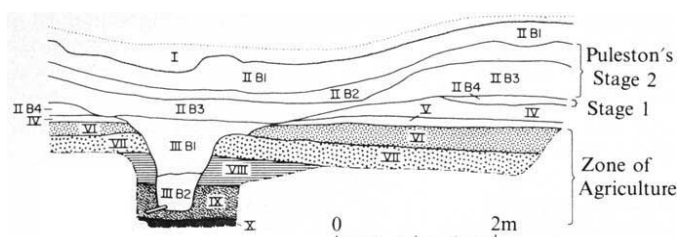


Fig. 2 Soil stratigraphy in Trench 2B, excavated by Puleston in 1973. Horizon designations based on Puleston's field notes and field and laboratory data of Bloom, Pohl and Buttleman for Pits 2X, 2Y and 2Z. Numerical designations do not imply chronological sequence of events.

Designation	Description	Colour*
I	Clay† topsoil, 4–8% o.c.	Brownish black (10YR 3/1)
II	Clay, o.c. <2% over the ancient fields, o.c. 4–10% in the ditches	Light grey (2.5Y 7/1) Brownish grey (10YR 5/1) Yellowish grey (2.5YR 4/1)
III	Shelly clay canal fill o.c. ≈ 4%	Brownish grey (10YR 4/1)
IV	Shelly marl, o.c. <1.5%	Light grey (2.5Y 7/1)
V	Shelly marl, o.c. ≈ 6%	Yellowish grey (2.5Y 5/1)
VI	High organic carbon clay o.c. ≈ 8%, many shells	Yellowish grey (2.5Y 4/1) with brownish black (2.5Y 3/1) charcoal fragments
VII	Clay with peaty bands o.c. ≈ 8%, few shells	Yellowish grey (2.5Y 6/1) and (2.5Y 4/1) with black (2.5Y 2/1) fragments in the bands
VIII	Banded peat, o.c. ≈ 16%	Yellowish grey (2.5Y 6/1) and (2.5Y 4/1) with black (2.5Y 2/1) fragments in the bands
IX	Sapric peat, o.c. ≈ 20%	Black (10YR 2/1)
X	Gleyed clay	Grey (N 5/0)

o.c., Organic carbon.

* Munsell colours.

† Laboratory classification after removal of calcium carbonate and gypsum.

mollusc, pollen, plant macrofossil, and ceramic samples by excavating three 1 × 2 m pits (2X, 2Y, 2Z in Fig. 1) to a depth of 2 m and by taking deeper samples with a bucket auger. We delimited the lateral extent of deposits by probing to a depth of 2 m at 8-m intervals along a 75-m transect extending from the river's edge to the bedrock outcrop. This report presents the results of our analyses unless specifically attributed to Puleston.

Soil stratigraphical units were differentiated on the basis of colour, morphology, content of organic matter and calcium carbonate (Figs 2 and 3). The stratigraphical units recorded numerous events: deposition of a noncultural substrate of clay, a period of agricultural activity and peat accumulation interrupted by sporadic natural sedimentation, and finally burial of the field surface under 1.5 m of clay poor in organic matter. The lowest unit (X) is a gleyed clay, a stratum found underlying most of the coastal strip and lower reaches of riverine environments in Belize³.

The peat units (VIII and IX) and the organic clay units (VI–VII) represent agricultural activity. Fragments of what we identified as carbonized stems of maize (*Zea*), collected by Puleston in 2J and by us in 2X, comprised as much as 68% of the botanical sample. Our analysis of pollen from 2X revealed up to 2% maize. The data suggest that in wetlands the Maya concentrated on corn production.

The high content of organic carbon in the soil indicating a poorly drained, non-oxidizing environment and the presence of freshwater molluscs (*Cochliopina*, *Pyrgophorus* and *Pomacea*) demonstrate that during this period the habitat was very wet or flooded much of the time. We believe that the Maya could only have cropped in the dry season (today, February–May) when the water table is lowest. Alternating bands of low organic matter, light coloured sediment found

within units VII and VIII demonstrate that even while the Maya were using their fields, sediment was accumulating, probably during high-water stages of the wet season.

The percentage of tree pollen (*Quercus*, *Terminalia*-type, *Moraceae* each <5%) was the lowest that has been found in an archaeological context in northern Belize. This is a probable indication of extensive forest clearance and suggests that the Maya used wetland agriculture in conjunction with wet-season swidden on uplands like many lowland cultivators today⁴.

Although no ceramic fragments occurred in units VI to IX, we have radiocarbon confirmation of Puleston's evidence that Maya cultivators were active on the Rio Hondo floodplain by the Middle Preclassic period. We obtained an uncorrected date of $2,626 \pm 190$ radiocarbon yr: 670 BC (DIC 2120) on the maize charcoal from unit IX in 2J. If the carbon isotope fractionation in maize is considered, the date may be 200 yr older. This determination is consistent with Puleston's date of $3,010 \pm 230$ yr BP on a cut post found lying diagonally penetrating units VII–IX in 2J.

The profile indicates that after initial agricultural activity, ditches were dug in the fields to encourage drainage. We suggest that the additional water control measures were necessary because the sea level was rising. High³ records a Holocene marine transgression that flooded much of the coast and fluvial systems of northern Belize. This transgression would have resulted in the drowning of the lower portion of the Rio Hondo. When the base level at the mouth of a river is raised, the river responds by depositing sediment, thus lowering its gradient. The Maya countered this rise in water level by ditching their fields.

Eventually, the continued transgression and aggradation resulted in the termination of agriculture. Units IV and V indicate shallow-water deposition without human intervention. The units are composed of shell marl that is about 70% calcium carbonate (Fig. 3). The distribution and relative densities of molluscan species (*Cochliopina* > *Pyrgophorus* > *Biomphalaria* > *Pomacea* > *Stenophysa* > *Gundlachia*) indicate permanent shallow water (littoral zone) over a period of several hundred years with possible fluctuations in water level resulting from river flooding. Unit III, a shelly clay canal fill, is contemporaneous with units IV and V.

Unit II is a complex unit of very fine-grained grey clay containing a few poorly preserved, vertically mixed ceramic fragments. Puleston originally postulated two stages of artificial platform building corresponding roughly to our units II B4 (his Stage 1) and II B2 and II B3 (his Stage 2). He thought that the Maya constructed Stage 1 out of floodplain sediments and

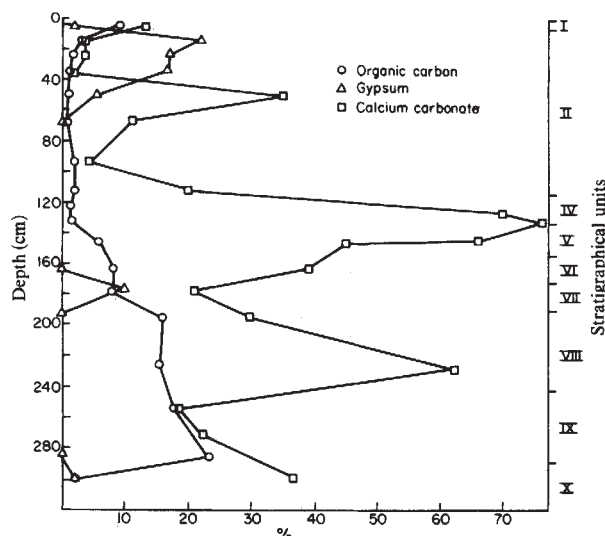


Fig. 3 Organic carbon, gypsum and calcium carbonate in Pit 2X sampled by Bloom, Pohl and Buttleman in 1980.

that they built up Stage 2 with poorly consolidated marl-like sediments brought from upland quarries.

Our research did not reveal conclusive evidence that the prehistoric Maya returned to farm at this location⁵. The ceramics, which in our sample ranged from terminal Preclassic to terminal late Classic types, might be attributed to the presence of settlement in the area rather than to the activity of farmers as Puleston thought. The concentrations of macrofossils and pollen of economic plants characteristic of units VI–IX were lacking. Units II–V, which were very low in organic N, total P and extractable K, were infertile compared with the soil of units VI–IX. The calcium carbonate peak at the 60-cm depth (Fig. 3) corresponds to a zone of mollusc deposition with a relative abundance of species similar to that found in units IV and V. This suggests that this zone of unit II was also deposited under permanent shallow water. The gypsum distribution is consistent with natural deposition. Our analysis of Rio Hondo water showed that it is nearly saturated with respect to gypsum. In the dry season when evapotranspiration exceeds precipitation, gypsic ground waters move upward through the soil profile depositing gypsum.

Funding was provided by the Title XII Strengthening Grant and the International Agriculture Programs of the College of Agriculture, University of Minnesota, by Florida State University and by the Wenner-Gren Foundation for Anthropological Research, Inc. This paper was written while M.P. was a Tinker Foundation postdoctoral fellow.

Received 25 October; accepted 18 November 1982.

1. Puleston, D. E. in *Social Process in Maya Prehistory* (ed. Hammond, N.) 449–467 (Academic, New York, 1977).
2. Wright, A. C. S. et al. *Land in British Honduras* Colonial Research Publ. 24 (1959).
3. High, L. R. Jr in *Carbonate Sediments, Clastic Sediments and Ecology* (eds Wantland, K. W. & Pusey, III, W. C.) 53–96 (American Association of Petroleum Geologists, Tulsa, 1975).
4. Wilk, R. R. thesis, Univ. Arizona (1981).
5. Antoine, P. P., Skarie, R. L. & Bloom, P. R. in *Maya Subsistence* (ed. Flannery, K. V.) 227–236 (Academic, New York, 1982).

Categorizing sounds and learning to read—a causal connection

L. Bradley & P. E. Bryant

Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, UK

Children who are backward in reading are strikingly insensitive to rhyme and alliteration¹. They are at a disadvantage when categorizing words on the basis of common sounds even in comparison with younger children who read no better than they do. Categorizing words in this way involves attending to their constituent sounds, and so does learning to use the alphabet in reading and spelling. Thus the experiences which a child has with rhyme before he goes to school might have a considerable effect on his success later on in learning to read and to write. We now report the results of a large scale project which support this hypothesis.

Our study combined two different methods. The first was longitudinal. We measured 403 children's skills at sound categorization before they had started to read, and related these to their progress in reading and spelling over the next 4 yr: at the end of this time the size of our group was 368. The second was intensive training in sound categorization or other forms of categorization given to a subsample of our larger group. We used both methods because we reasoned that neither on its own is a sufficient test of a causal hypothesis and that the strengths and weaknesses of the two are complementary. Properly controlled training studies demonstrate cause–effect relationships, but these could be arbitrary; one cannot be sure that such relationships exist in real life. On the other hand

longitudinal studies which control for other variables such as intelligence do demonstrate genuine relationships; but it is not certain that these are causal. For example simply to show that children's skills at categorizing sounds predict their success in reading later on would not exclude the possibility that both are determined by some unknown *tertium quid*. Thus the strength of each method makes up for the weakness of the other. Together they can isolate existing relationships and establish whether these are causal.

This combination of methods has not been used in studies of reading or, as far as we can establish, in developmental research in general.

Initially we tested 118 4-yr-olds and 285 5-yr-old children (Table 1) on categorizing sounds. None of the children could read (that is, were able to read any word in the Schonell reading test). Our method, as before¹ was to say three or four words per trial, all but one of which shared a common phoneme (Table 2): the child had to detect the odd word. There were 30 trials. In such a task the child must remember the words as well as categorize their sounds. To control for this we also gave them 30 memory trials: the child heard the same words and had to recall them straightaway. In addition we tested verbal intelligence (EPVT).

At the end of the project (as well as at other times) we gave the children standardized tests of reading and spelling, and we also tested their IQ (WISC/R) to exclude the effects of intellectual differences. To check that our results were specific to reading and spelling and not to educational achievement in general we also included a standardized mathematical test (MATB-NFER), which we administered to 263 of our total sample of 368.

There were high correlations between the initial sound categorization scores and the children's reading and spelling over 3 yr later (Table 3). Stepwise multiple regressions established that these relationships remained strong even when the influence of intellectual level at the time of the initial and the final tests and of differences in memory were removed (Table 3). In every case categorizing sound accounted for a significant proportion of the variance in reading and spelling with these other factors controlled.

So a definite relationship does exist between a child's skill in categorizing sounds and his eventual success in reading and spelling. The design of the project, for the reasons just given, included a training study as a check that any such relationship is a causal one. 65 children were selected from our sample and divided into four groups closely matched for age, verbal intelligence and their original scores on sound categorization. These children were drawn from those with lower scores on sound categorization (at least two standard deviations below the mean); they could not read when the training began. Starting in the second year of the project two of the groups (I and II) received intensive training in categorizing sounds. The training involved 40 individual sessions which were spread over 2 yr.

Table 1 Details of sample

	Children initially tested at age 4 yr	Children initially tested at age 5 yr
<i>N</i> at end of project	104	264
<i>Initial tests</i>		
Mean age (months)	58.62	65.52
Mean EPVT	110.62	109.39
<i>Final tests</i>		
Mean age (months)	101.85	101.42
Mean IQ (WISC)	113.38	106.79
Mean reading age (months)		
Schonell	103.13	100.03
Neale	105.13	101.30
Mean spelling age (months)		
Schonell	97.27	93.94

Table 2 Examples of words used in initial sound categorization tests and mean scores on these tests

	4-yr group			5-yr group			
	Words given to children			Words given to children			
Sounds in common							
First sound	hill	pig	pin	bud	bun	bus	rug
	bus	bun	rug	pip	pin	hill	pig
Middle sound	cot	pot	hat	lot	cot	hat	pot
	pin	bun	gun	fun	pin	bun	gun
End sound	pin	win	sit	pin	win	sit	fin
	doll	hop	top	doll	hop	top	pop

Standard deviations given in parentheses.

With the help of coloured pictures of familiar objects the children were taught that the same word shared common beginning (hen, hat), middle (hen, pet) and end (hen, man) sounds with other words and thus could be categorized in different ways. Group I received this training only, but group II in addition was taught, with the help of plastic letters, how each common sound was represented by a letter of the alphabet (see ref. 2 for further details of this method). The other two groups were controls. Group III was also taught over the same period in as many sessions and with the same pictures how to categorize but here the categories were conceptual ones; the children were taught that the same word could be classified in several different ways (for example, hen, bat (animals); hen, pig (farm animals)). Group IV received no training at all.

The training had a considerable effect which was specific to reading and spelling (Table 4). At the end of the project group I (trained on sound categorization only) was ahead of group III (trained on conceptual categorization only) by 3–4 months in standardized tests of reading and spelling. This suggests a causal relationship between sound categorization and reading and spelling. Group II (trained with alphabetic letters as well as on sound categorization) succeeded even better than group I (trained on sound categorization only) in reading and particularly in spelling. This suggests that training in sound categorization is more effective when it also involves an explicit connection with the alphabet. That the relationship is specific to these two skills is shown by the mathematics results where the differences were a great deal smaller.

Analyses of covariance, in which the covariates were the children's final IQ scores and their age at the time of the final reading and spelling tests, established that the group differences were significant in the case of reading (Schonell: $F = 5.23$; d.f.3,58; $P < 0.003$. Neale: $F = 7.80$; d.f.3,58; $P < 0.001$) and of spelling ($F = 12.18$; d.f.3,58; $P < 0.001$) but not in the case of mathematics ($F = 1.64$; d.f.3,39; P , not significant). Post tests (Tukey's HSD) showed that group II was significantly better than both control groups (groups III and IV) in Schonell and in Neale reading ($P < 0.05$) and in Schonell spelling ($P < 0.01$). There was no significant difference between groups I and II

Table 3 Correlations between initial sound categorization and final reading and spelling levels

Final scores	Initial scores:					
	Sound categorization		EPVT		Memory	
	4	5	4	5	4	5
Reading: Schonell	0.57	0.44	0.52	0.39	0.40	0.22
Reading: Neale	0.53	0.48	0.52	0.44	0.40	0.25
Spelling: Schonell	0.48	0.44	0.33	0.31	0.33	0.20

Multiple regressions testing relationship of initial sound categorization to final reading and spelling levels

	Schonell reading		Neale reading		Schonell spelling	
	4	5	4	5	4	5
% Of total variance accounted for by all variables	47.98	29.88	47.55	34.52	33.59	24.77
% Of total variance accounted for by sound categorization*	9.84†	4.06†	6.24†	4.56†	8.09†	5.59†

* IQ, EPVT, final CA and memory controlled.

† $P < 0.001$.

(the two groups trained in sound categorization) in the two reading tests but group II did surpass group I in spelling ($P < 0.05$). Although reading and spelling scores in group I were always ahead of those of group III this difference did not reach significance in the post tests. But the consistent 3–4-month superiority of group I over group III does strongly suggest that training in sound categorization affects progress in reading and spelling. Group I was significantly better than group IV (the untrained control group) in the two reading tests and in the spelling test ($P < 0.05$). On the other hand there were no significant differences at all between the two control groups (III and IV).

Table 4 Training study: details of groups and mean final reading, spelling and mathematics levels

Groups	Mean scores			
	Experimental groups		Control groups	
N	I	II	III	IV
	13	13	26	13
Aptitude tests				
Initial EPVT	103.00	103.00	102.34	102.69
Final IQ (WISC/R)	97.15	101.23	102.96	100.15
Final educational tests				
Schonell: reading age (months)	92.23	96.96	88.48	84.46
Neale: reading age (months)	93.47	99.77	89.09	85.70
Schonell: spelling age (months)	85.97	98.81	81.76	75.15
N	9	8	20	7
Maths MATB (ratio score)	91.27	91.09	87.99	84.13

Reading, spelling and mathematics mean scores are adjusted for two covariates: age and IQ.

Put together our longitudinal and training results provide strong support for the hypothesis that the awareness of rhyme and alliteration which children acquire before they go to school, possibly as a result of their experiences at home, has a powerful influence on their eventual success in learning to read and to spell. Although others have suggested a link between phonological awareness and reading³⁻⁵ our study is the first adequate empirical evidence that the link is causal. Our results also show how specific experiences which a child has before he goes to school may affect his progress once he gets there.

We thank Morag Maclean for help with gathering and analysing the data, the Oxford Education Authority and the schools for their cooperation, and the SSRC for supporting our research.

Received 13 September; accepted 17 November 1982.

1. Bradley, L. & Bryant, P. E. *Nature* **271**, 746-747 (1978).
2. Bradley, L. *Assessing Reading Difficulties* (Macmillan, London, 1980).
3. Goldstein, D. M. *J. educ. Psychol.* **68**, 680-688 (1976).
4. Liberman, I. et al. in *Toward a Psychology of Reading* (eds Reber, A. & Scarborough, D.) (L. Erlbaum Association, Hillsdale, New Jersey, 1977).
5. Lundberg, L., Olofsson, A. & Wall, S., *Scand. J. Psychol.* **21**, 159-173 (1980).

Chloride impermeability in cystic fibrosis

Paul M. Quinton

* Division of Biomedical Sciences, University of California, Riverside, California 9252-0121, USA and Department of Physiology, UCLA Medical School, Los Angeles, California 90024, USA

Cystic fibrosis is the most common fatal genetic disease affecting caucasians and is perhaps best characterized as an exocrinopathy involving a disturbance in fluid and electrolyte transport¹. A high NaCl concentration in the sweat is characteristic of patients with this disease; the basic physiological reason for this abnormality is unknown. We have microperfused isolated sweat ducts from control subjects and cystic fibrosis patients, and report here results which suggest that abnormally low Cl⁻ permeability in cystic fibrosis leads to poor reabsorption of NaCl in the sweat duct, and hence to a high concentration of NaCl in the sweat.

Sweat glands were microdissected from skin biopsies taken from five control subjects and three cystic fibrosis patients. All subjects were males between 23 and 28 yr old. Single sweat glands were dissected further to isolate the reabsorptive duct from the secretory coil. One end of the reabsorptive duct was then cannulated with a micropipette² which also served as a microelectrode in the lumen of the duct. To avoid significant changes in the composition of the perfusion fluids due to reabsorptive activity, duct lengths of 1.5-2.0 mm were perfused at rates >100 nl min⁻¹ at 25 °C. The bathing solution always contained 150 mM NaCl while the duct was perfused with solutions of different composition (see Table 1). Transepithelial potential differences during perfusions were recorded via saturated KCl bridges connecting the bathing solution and the perfusion micropipette to a high impedance electrometer (Kiethley 610C) using Ag-Ag/Cl electrodes. Asymmetries without the duct in place were usually <±1.0 mV and were subtracted from all measurements.

During periods of perfusion in which there were identical concentrations (150 mM) of NaCl in the bath and in the lumen, the average spontaneous potential in control ducts was -6.8 mV, with the lumen negative with respect to the bath, while in the same conditions the average potential in ducts from cystic fibrosis patients was -76.9. This markedly increased negative potential in the cystic fibrosis duct cannot be due to increased Na⁺ transport as suggested for cystic fibrosis respira-

Table 1 Average luminal potential differences in cystic fibrosis and normal sweat ducts perfused with solutions of varying composition

	150 mM NaCl	50 mM NaCl	75 mM Na ₂ SO ₄
Pre-ouabain control (n = 7)	-6.8 ± 2.5	-24.2 ± 2.4	-75.5 ± 11.1
CF (n = 5)	-76.9 ± 13.2*	-47.6 ± 12.3†	-33.3 ± 19.5
Post-ouabain control (n = 7)	-0.9 ± 0.5	-14.8 ± 1.3	-42.3 ± 6.5
CF (n = 5)	-0.9 ± 1.7	+9.9 ± 3.8*	-4.4 ± 1.7†

Values shown are the average luminal potential differences in sweat ducts microperfused with solutions of different composition before and after exposing the contraluminal surface of the duct to 5×10^{-6} M ouabain. In addition to the principal solute indicated, all solutions contained: 2.5 mM KCl, 2.5 mM NaH₂PO₄, 1.0 mM CaCl₂, 1.0 mM MgSO₄ and 10.0 mM glucose. Sufficient mannitol was added to hypotonic solutions to make all solutions iso-osmotic with plasma. More than one sweat duct from two of the control and two of the cystic (CF) subjects were microperfused so that the results represent the mean ± s.e.m. of seven control sweat ducts and five cystic fibrosis sweat ducts. Significant differences were tested by Student's *t*-test.

**P* < 0.001; †*P* < 0.05.

tory tissues³, as ductal reabsorption of Na⁺ is only about one-third to one-fifth of that in normal glands^{4,5}. It therefore seems likely that the elevated potential is due to a separation of charge brought about by a difference in the permeability of anions in the two tissues. To test this possibility, a 1:3 dilution diffusion potential for NaCl (dilute solution in the lumen) and a bi-ionic diffusion potential for Na₂SO₄ (lumen) and NaCl (bath) were measured before and after inhibiting electrolyte transport with ouabain. In control ducts, the potential in the lumen before ouabain rose to -24.2 mV when the perfusate was 50 mM NaCl, indicating that Cl⁻ is more permeable than Na⁺ in control ducts. Changing the perfusate to Na₂SO₄ in which the SO₄ anion is assumed to be relatively impermeable, caused the potential to increase further to -75.5 mV, which is almost identical to the potential observed in cystic fibrosis ducts perfused with 150 mM NaCl (Table 1). This result argues strongly that the mechanisms for active Na⁺ reabsorption in the cystic fibrosis ducts are not impaired and that Cl⁻ passively follows Na⁺ out of the normal duct. In the cystic fibrosis duct, the luminal potential fell to -47.6 mV with 50 mM NaCl as the perfusate and to -33.3 mV with Na₂SO₄ as the perfusate (Table 1). The drop in the potential with lower concentrations of NaCl in the lumen may be due to some extent to the diffusion potential contribution, but the reason for the decrease during perfusion with Na₂SO₄ is not clear. As the potential in one duct did not fall during this perfusion, the decrease may be related to a variable, time-related reduction in tissue activity *in vitro*.

Treating the serosal side of the tissue with ouabain virtually abolished the spontaneous potential difference (Table 1). Nonetheless, the average dilution diffusion potential in control ducts was -14.8 mV, but in cystic fibrosis ducts the polarity of the potential was reversed with an average value of +9.9 mV. The change in sign for cystic fibrosis ducts in the absence of active transport, clearly demonstrates that the passive permeability properties of the epithelium are significantly different and reversed compared with normal ducts. Using the 'constant-field equation' (refs 6, 7) to solve for the ratio of the Na⁺ permeability to the Cl⁻ permeability (P_{Na}/P_{Cl}) from the potentials generated in each tissue, we found that P_{Na}/P_{Cl} for controls was 0.26 while for cystic fibrosis patients it was 2.3. Thus the Cl⁻ permeability in control ducts is almost an order of magnitude higher than in cystic fibrosis ducts. This calculated difference may be significantly underestimated as measured potentials may be less

* Address for correspondence.

than actual potentials due to the possibility of tissue trauma. Calculations using the maximum potentials measured, rather than the average showed that Cl^- permeability in control ducts may be more than 40 times that in cystic fibrosis ducts.

Although we have not yet measured absolute permeabilities of Cl^- , the small potentials generated during perfusion with Na_2SO_4 after ouabain treatment show that Cl^- permeability in the cystic fibrosis duct must be very low. The average bi-ionic potential generated across control ducts after ouabain was -42.3 mV, indicating that Cl^- is significantly more permeable than SO_4 . In contrast, this potential in cystic fibrosis ducts was only -4.4 mV, indicating that Cl^- is only slightly more permeable than SO_4 . As SO_4 is generally considered to be almost impermeable across epithelia, the relatively low bi-ionic potential with Cl^- in cystic fibrosis ducts strongly suggests that Cl^- also has a very low permeability in this genetic variant of the tissue.

We believe that these data provide strong evidence that abnormally low Cl^- permeability is the basis of poor reabsorption of NaCl in the cystic fibrosis sweat duct and the correspond-

ing high concentration of NaCl in the sweat of these patients. Previously, it was reported that the potential difference across the respiratory epithelium in cystic fibrosis patients was dramatically increased compared with controls³, which was interpreted as being due to increased rates of Na^+ reabsorption in the airways. In view of our findings, and as this is a genetic disease which should be expressed as a common defect in affected tissues, we believe that the observed potentials in the airways are more likely to be due to an abnormally low Cl^- permeability than to increased Na^+ transport. If this is the case, it seems possible that a generalized defect in the Cl^- permeability may be closely associated with the fundamental disturbance in this disease and may further elucidate the problems associated with characteristic abnormalities of the pancreas, intestine and lung.

This work was supported by awards from the Getty Oil Company, the Gillette Co., and the USPHS-NIH grant no. AM 26547-05. I thank Ms Diane Hannon and Ms Rebecca Laude for technical assistance and Dr Richard Dooley for assistance with patients.

Received 28 September; accepted 23 November 1982.

1. Mangos, J. A. in *Physiology of Membrane Disorders* (eds Andreoli, T. E., Hoffman, J. F. & Fanestil, D. D.) 941-954 (Plenum, New York, 1978).
2. Quinton, P. M. *Pflügers Arch. ges. Physiol.* **391**, 309-313 (1981).
3. Knowles, M., Gatzky, T. & Boucher, R. *New Engl. J. Med.* **305**, 1489-1495 (1981).

4. Schulz, I. J. *J. clin. Invest.* **48**, 1470-1477 (1969).

5. Quinton, P. M. in *Fluid and Electrolyte Abnormalities in Exocrine Glands in Cystic Fibrosis* (eds Quinton, P. M., Martinez, J. R. & Hopper, U.) (San Francisco Press, 1982).

6. Goldman, D. E. *J. gen. Physiol.* **27**, 37-60 (1943).

7. Hodgkin, A. L. & Katz, B. *J. Physiol. Lond.* **108**, 37-77 (1949).

T^{Or1} is a novel, variant form of mouse chromosome 17 with a deletion in a partial t haplotype

Lee M. Silver, Deborah Lukralle & James I. Garrels

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

Moutier¹ discovered, in a mouse from a noninbred Swiss/Orleans laboratory stock, a spontaneous dominant mutation which mapped to the T locus, and which was named T^{Or1} . Genetic analyses indicated that T^{Or1} was not a simple mutation at one locus, but rather a deletion over a 3-centimorgan region of chromosome 17 that included both T and *quaking* (*qk*)^{2,3}. Further experiments reported by Erickson *et al.*³, and a more comprehensive study by Hammerberg⁴, have demonstrated that T^{Or1} is associated with recessive genetic properties affecting sperm function, characteristic of the proximal region of complete t haplotypes. These results were interpreted as evidence for the location of proximal t haplotype 'sperm factors' within the region deleted by T^{Or1} . We now provide conclusive biochemical and genetic evidence that the ' T^{Or1} haplotype' is inseparably associated with a chromosomal region derived from a naturally occurring mouse t haplotype. Hence, it is likely that the t haplotype properties of T^{Or1} are a consequence not of the deletion itself, but of closely linked mutant t haplotype genes.

The dominant mutation *Brachyury* (T) on a mouse chromosome 17 causes a reduction in the tail length of $T/+$ heterozygotes, and embryonic lethality in T/T homozygotes⁵. The T mutation also interacts with naturally occurring, recessive t haplotypes to cause taillessness in compound T/t heterozygotes. This last property of T has been used to identify cryptic t haplotypes carried by animals from many wild mouse populations. The new mutation T^{Or1} was found to exhibit all of the classical phenotypic properties of the original T mutation: $T^{\text{Or1}}/+$ heterozygotes have a short tail; both T^{Or1}/T and $T^{\text{Or1}}/T^{\text{Or1}}$ embryos die *in utero*^{3,6}; and T^{Or1}/t compound heterozygotes are tailless. In addition, the T^{Or1} mutation acts as a recessive mutant allele at the *qk* locus—all T^{Or1}/qk animals

express the recessive quaking phenotype of *qk* homozygotes^{2,3}, and T^{Or1}/qk males also express the sterility phenotype characteristic of *qk/qk* males³. The accumulated data indicate that the T^{Or1} mutation is actually a deletion spanning at least the 3-centimorgan region of DNA between the loci of T and *qk* on chromosome 17.

The *Tcp-1* gene is located within the t complex and encodes a major testicular cell protein (TCP) called p63/6.9 (ref. 7). All naturally occurring t haplotypes carry the allele *Tcp-1*^a, which encodes a variant protein p63/6.9a. With one exception, all of the hundreds of examples of chromosome 17 not obviously carrying a t haplotype analysed to date from inbred and feral mice express a different form of the protein, p63/6.9b (ref. 8). The single exception is that of T^{Or1} , which does express p63/6.9a (ref. 9), suggesting that the T^{Or1} mutation is, indeed, associated with a short segment of t haplotype DNA¹⁰.

A higher resolution, comparative two-dimensional gel analysis of TCPs synthesized by inbred mice congenic for a series

Table 1 Transmission ratios of males carrying the T^{Or1} haplotype in *trans* to a distal t haplotype

Male	Distal t haplotype	Offspring			
		T^{Or1}	t^{distal}	% t^{distal}	χ^2
1715	t^{h17}	1	7	88	4.50
2083	t^{h17}	2	21	91	15.69
2130	t^{h17}	0	29	100	29.00
2701	t^{h17}	1	10	91	7.36
3167	t^{h17}	1	19	95	16.20
2568	t^{Lub2}	2	50	96	44.30
2708	t^{Lub2}	2	24	92	18.61
Total	t^{h17} and t^{Lub2}	9	160	95	134.92

The matings shown schematically in Fig. 3 were set up to obtain the genotypes used in this experiment: $T^{\text{Or1}}/t^{\text{h17}}$ and $T^{\text{Or1}}/t^{\text{Lub2}}$ (see also Fig. 1 for a diagrammatic representation of these genotypes). Each male was mated with outbred non- t haplotype (+/+) females. Progeny that receive the T^{Or1} chromosome have a short tail, whilst progeny that receive the t^{h17} or t^{Lub2} chromosomes have a normal tail. These phenotypes were scored at birth. χ^2 values were calculated in comparison with the expected mendelian values.

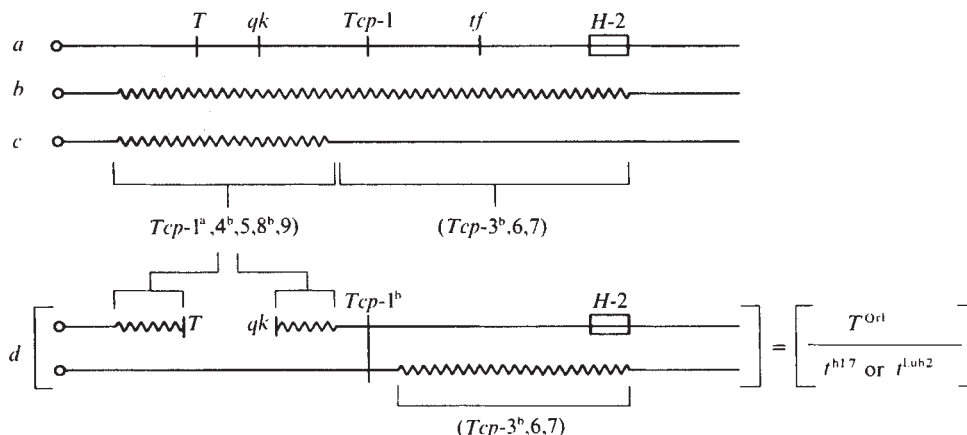


Fig. 1 Genetic map of *t* complex proteins (*Tcp*) genes, and a schematic representation of the Tl^{Or1}/t^{distal} genotypes used in the transmission ratio experiment. *a*, The *t* complex region of chromosome 17 with the markers: *T*, *Brachyury*; *qk*, *quaking*; *tf*, *tufted*; *Tcp-1*, *t* complex protein 1; and *H-2*, the major histocompatibility complex. *b*, A complete *t* haplotype spans the region indicated by the zigzag line. *c*, A representative proximal, partial *t* haplotype extends only over a portion of the *t* complex. Five *Tcp* genes have been mapped to the proximal region, and three *Tcp* genes to the distal region¹¹. *d*, A schematic representation of the Tl^{Or1}/t^{h17} and Tl^{Or1}/t^{Luh2} genotypes.

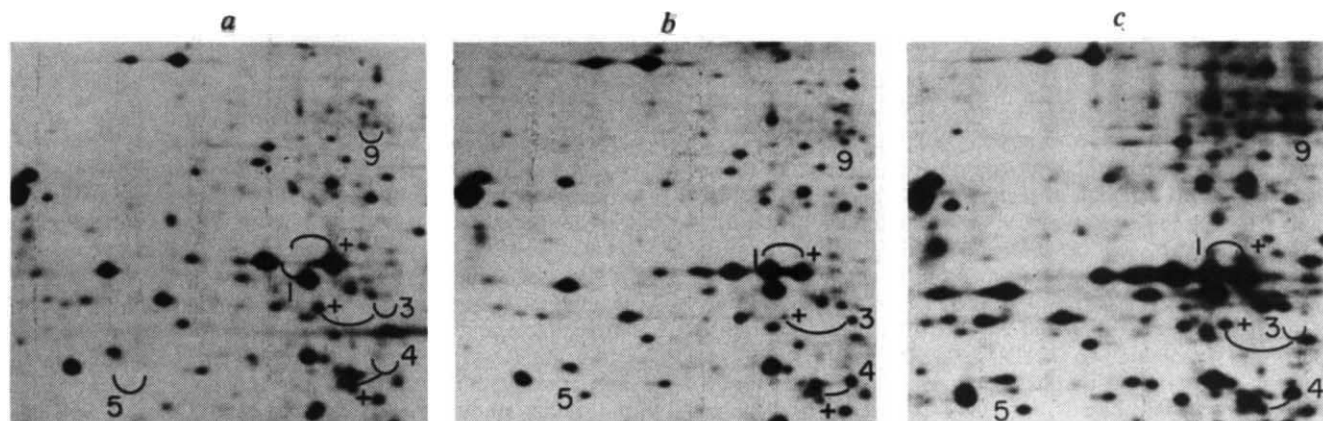


Fig. 2 Two-dimensional testicular cell protein patterns from mice carrying a wild-type chromosome 17, a complete *t* haplotype, or the Tl^{Or1} chromosome. The genotypes are: *a*, 129/SvJ (+/+); *b*, 129-(+/t^{w5})N17; *c*, 129- Tl^{Or1} F1. Sample preparation and gel electrophoresis was performed as described previously^{7,11,17}. Between approximately 0.1 and 0.25 μ Ci were loaded onto each gel in a volume of 10 μ l. The first dimension isoelectric focusing gels contained pH 6–8 ampholytes (LKB), and the second dimension SDS gels contained 12.5% acrylamide. Only a small segment of the complete gel pattern is shown in each frame of this figure. The acidic side of each protein pattern is on the left. The Tl^{Or1} haplotype expresses the *t* haplotype-specific proteins p63/6.9a (1), TCP-4B (4), TCP-5, TCP-9 and TCP-8B (not present in this portion of the gel pattern). The wild-type analogues of p63, TCP-3 and TCP-4 are indicated by connecting lines and the (+) symbol. The absence of a particular *t* haplotype-specific protein is indicated by an empty semicircle.

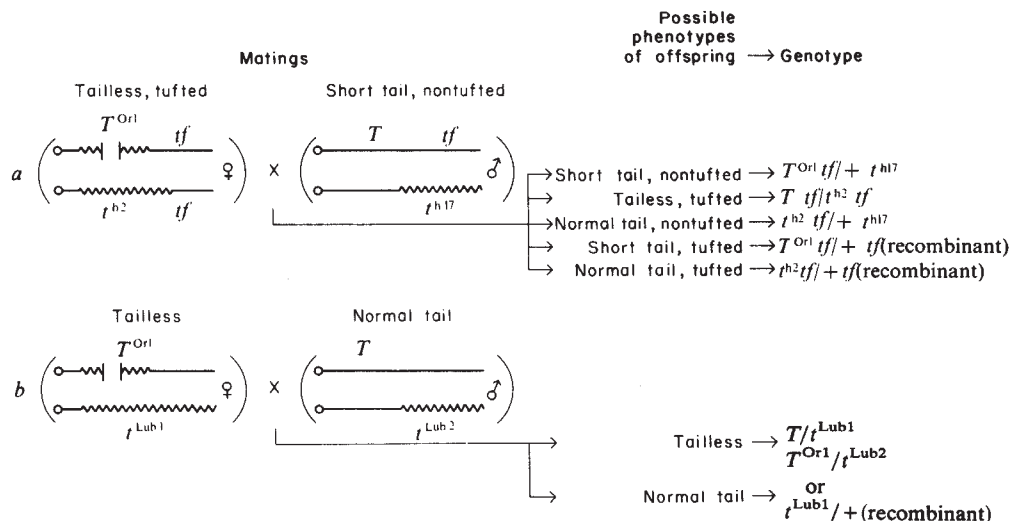
of lethal *t* haplotypes has allowed the identification of seven additional *t* haplotype-specific proteins, TCP-3 to TCP-9¹¹. The eight *t* haplotype-specific proteins are expressed in common by all complete *t* haplotypes studied. Alternative, wild-type forms for five of these TCPs have been identified, and evidence from cell-free translation experiments suggests that these proteins (p63/6.9, TCP-3, TCP-4, TCP-7, TCP-8) and TCP-5 are encoded directly by genes within the *t* complex. Partial *t* haplotypes have been used to map the *Tcp* genes to proximal or distal portions of the *t* complex, as shown in Fig. 1.

To investigate further the possibility that the Tl^{Or1} deletion is associated with a region of mutant *t* chromatin, a high resolution, two-dimensional gel analysis was performed on ³⁵S-methionine labelled TCPs from animals heterozygous for the Tl^{Or1} mutation. The results of this experiment demonstrate that all five of the proximally located *t* haplotype-specific proteins are expressed by animals carrying Tl^{Or1} (see Fig. 2). Our previous study¹¹ suggests that at least four of these proteins—p63/6.9, TCP-4, TCP-5 and TCP-8—are independent,

unmodified gene products encoded by different mRNAs. The result presented here clearly demonstrates the association of the Tl^{Or1} mutation with an extended segment of *t* haplotype DNA carrying a series of genes expressed in the testes. We suggest that this mutant complex—part deletion and part *t* haplotype—be renamed the Tl^{Or1} haplotype. A possible explanation for the origin of this bizarre chromosome has been discussed previously¹⁰.

Complete *t* haplotypes carry two separable but interactive factors that are involved in causing a reduction in male fertility^{4,12,13}. Males that carry two complementing, complete *t* haplotypes are unconditionally sterile, while males carrying a complete *t* haplotype opposite an isolated proximal or distal factor are either completely sterile or quasi-sterile. It has been clearly demonstrated^{3,4} that the Tl^{Or1} haplotype has all of the properties expected of an isolated proximal *t* haplotype sterility factor (which I have called *tcs-1*)¹⁰. When males carry Tl^{Or1} opposite a complete *t* haplotype (t^0 , t^{12} , or t^{w2}), they exhibit greatly reduced fertility. We have confirmed this result in the case of Tl^{Or1}/t^{Luh1} males (data not shown).

Fig. 3. Mating schemes set up for the construction of compound heterozygotes carrying T^{Or1} opposite either t^{h17} or t^{Lub2} . All short tailed, nontufted animals obtained as offspring from mating scheme *a* must carry T^{Or1} opposite t^{h17} , because the wild-type allele of the tufted locus is inseparable from the t^{h17} haplotype. The distal haplotypes t^{h17} and t^{Lub2} were derived by recombinant events from t^6 and t^{Lub1} respectively. t^{h17} does not carry the t^6 tail interaction factor, and the tail phenotype of T/t^{h17} animals is identical to that of $T/+$ animals. However, t^{Lub2} manifests an altered interaction with the T locus, such that T/t^{Lub2} animals have tails of normal length¹⁵. This enhancement of tail length or 'suppression' of the mutant T tail effect has also been observed by Lyon for t^{h7} (ref. 17), and by Klein for g^{Tu3} (offspring of mating scheme *b* must have a non- t haplotype chromosome recessive can never produce short-tailed offspring have a T^{Or1}/t^{Lub2} genotype).



At least three separable factors (one proximal, one central and one distal) interact to distort male transmission ratio¹⁴. Only a complete *t* haplotype can be transmitted consistently at a very high ratio—loss of either the proximal end or the distal end causes a loss of the high transmission. To determine whether the Tt^{Or1} haplotype carries a proximal *t* haplotype distortion factor (which I have called *Tcd-1*)¹⁰, we set up a series of crosses (see Fig. 3) to obtain two parallel genotypes in which Tt^{Or1} was present in *trans* to a partial, distal *t* haplotype (see Fig. 1). Each of the distal *t* haplotypes used (t^{Lub2} or t^{h17}) was derived originally by recombination from a complete *t* haplotype with the resultant loss of the *Tcd-1* factor^{14,15}, and several *Tcp* genes (unpublished observations). Hence, isolated t^{Lub2} and t^{h17} chromosomes are not transmitted from males at high ratios¹⁴, and neither is Tt^{Or1} (ref. 4). Tt^{Or1} , t^{h17} , and t^{Lub2} each carry only a subset of the eight variant *Tcp* genes associated with all complete *t* haplotype as shown in Fig. 1. However, a genotype with Tt^{Or1} in combination with either t^{h17} or t^{Lub2} reconstitutes the complete set of eight variant *Tcp* genes. If the three *t* haplotype distortion factors can interact in *trans* (as indicated by the results of Styryna and Klein¹⁶), and if the Tt^{Or1} haplotype carries the proximal factor *Tcd-1*, then Tt^{Or1}/t^{h17} and Tt^{Or1}/t^{Lub2} males should exhibit severe transmission ratio distortion.

The results from this experiment dramatically demonstrate that the Ti^{Or1} haplotype can function in *trans* to promote a high transmission frequency for each of the partial, distal t haplotypes studied (Table 1). The cumulative transmission ratio obtained for the 169 offspring born to 7 males was 95%. As, neither Ti^{Or1} nor the distal t haplotypes are transmitted at a high frequency when isolated in separate genotypes, it seems likely that the synergistic effect reported here for *trans* heterozygotes is a consequence of a positive interaction among gene products from Ti^{Or1} and a distal t haplotype. This implies that it is the t chromatin associated with the Ti^{Or1} haplotype and not the deletion that promotes the high transmission frequency. In support of this view, the T^{hp} deletion also extends over T and qk , but does not exhibit any t effects on fertility or transmission ratio (ref. 7 and unpublished observations). Hence, the proximal t haplotype sperm factors must be independent of the T locus and map either proximal to T or distal to qk within the t chromatin associated with Ti^{Or1} (as shown in Fig. 1). It is possible that one or two of the t -specific testicular cell

proteins encoded by *TfOr1* is a primary gene product of the proximal *t* sperm factors; however, further biochemical studies will be necessary to confirm any gene product-phenotype associations.

One last point is the determination of which partial *t* haplotype becomes highly transmitted from compound heterozygotes carrying proximal and distal haplotypes in *trans*. While the distal haplotypes were highly transmitted in the present study, the proximal haplotypes were highly transmitted in a similar study reported by Styrna and Klein¹⁵. These contrasting results can be understood in the context of the pivotal role of the centrally located distortion factor called A by Lyon and Mason¹⁴. According to Lyon¹², the A factor determines the chromosome to be distorted in a complex genotype. Thus, the distal *t* haplotypes studied herein would appear to carry the central A factor, while the proximal *t* haplotypes studied by Styrna and Klein would carry the A factor instead. It is possible, of course, that the situation is more complex than it appears to be and that there are more than three interacting factors that actually mediate transmission ratio distortion in the *t* complex.

We thank Jane Uman, Jane Emanuele and Lynn Cascio for technical assistance. This research was supported by NIH grants HD16091 to L.M.S. and CA13106 to J.I.G.

Received 4 November; accepted 6 December 1982.

1. Moutier, R. *Mouse News Lett.* **48**, 38 (1973).
2. Moutier, R. *Mouse News Lett.* **49**, 42 (1973).
3. Erickson, R. P., Lewis, S. E. & Slusser, K. S. *Nature* **274**, 163-164 (1978).
4. Hammerberg, C. *Genet. Res.* **37**, 71-77 (1981).
5. Chesley, P. J. *exp. Zool.* **70**, 429-459 (1935).
6. Bennett, D. *et al. Genet. Res.* **26**, 95-108 (1975).
7. Silver, L. M., Artzt, K. & Bennett, D. *Cell* **17**, 275-284 (1979).
8. Silver, L. M., White, M. & Artzt, K. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6077-6080 (1980).
9. Altan, A., Silver, L. M., Artzt, K. & Bennett, D. *Nature* **288**, 368-370 (1980).
10. Silver, L. M. *Genet. Res.* **38**, 115-123 (1981).
11. Silver, L. M., Uman, J. & Garrels, J. I. *Cell* (submitted).
12. Lyon, M. F. *Symp. zool. Soc. Lond.* **47**, 455-477 (1981).
13. Bennett, D. & Dunn, L. C. in *Proc. Symp. Immunogenetics of the H-2 System* (ed. Karger, S.) 90-103 (Liblice-Prague, Basel, 1971).
14. Lyon, M. F. & Mason, I. *Genet. Res.* **29**, 255-266 (1977).
15. Winking, H. *Mouse News Lett.* **59**, 33 (1978).
16. Styma, J. & Klein, J. *Genet. Res.* **38**, 315-325 (1981).
17. Lyon, M. F. & Meredith, R. *Heredity* **19**, 313-325 (1964).

Aberrant rearrangement of the κ light-chain locus involving the heavy-chain locus and chromosome 15 in a mouse plasmacytoma

Brian G. VanNess*§, Marjorie Shapiro*, Dawn E. Kelley*, Robert P. Perry*, Martin Weigert*, Peter D'Eustachio† & Frank Ruddle‡

* The Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111, USA

† Department of Biochemistry, New York University, New York 10016, USA

‡ Department of Biology, Yale University, New Haven, Connecticut 06511, USA

The creation of a functional antibody gene requires the precise recombination of gene segments initially separated on the chromosome. Frequently errors occur in the process, resulting in the formation of a non-functional gene. The non-functional genes can be generated by incomplete rearrangements¹, frame-shifts², or the use of pseudo V or J joining segments³. It is likely that these aberrant rearrangements arise by the same mechanism as is used in generating functional genes, a process which we have suggested may involve unequal sister chromatid exchange⁴. Aberrant rearrangements of immunoglobulin genes occur in normal lymphocytes and play a major part in allelic exclusion⁵. However, it has recently been suggested that aberrant rearrangements involving immunoglobulin and non-immunoglobulin genes may be involved in tumorigenesis. This suggestion has been stimulated by the frequent occurrence of translocations involving chromosomes known to carry immunoglobulin genes in B-cell malignancies. The rearrangement of non-immunoglobulin DNA to the heavy-chain locus has recently been reported^{6,7}. Some aberrant rearrangements of the κ locus appear to be due to rearrangements to sites that do not include the conventional sequence for V gene segment joining⁸. Here we describe an aberrant κ rearrangement that has led to the joining of DNA from chromosomes 15, 6 and 12, and so appears to be the result of chromosomal translocations or transpositions. As 15/6 or 15/12 translocations have frequently been found in mouse plasmacytomas (as have analogous translocations in human lymphocyte tumours) this aberrant κ rearrangement may be unique to the plasmacytoma from which it was isolated.

Surveys of κ rearrangements in κ -expressing mouse plasmacytomas have shown that approximately 50% have one productively rearranged (κ^+) and one unrearranged (κ^0) allele (κ^+/κ^0 cells), while the other 50% have rearranged both alleles^{9,10}, one productively and one (κ^-) aberrantly (κ^+/κ^- cells). This frequency of aberrant rearrangements is similar to, but significantly higher than, the natural frequency (40% κ^+/κ^- cells) in normal κ -expressing splenic B-cells⁵. Therefore, tumour specific rearrangements might contribute to the κ^- frequency in the plasmacytoma survey. One aberrant rearrangement that appears to be of this type is found in the NZB plasmacytoma, PC7183 (ref. 4 and Fig. 1). By using a series of κ probes against Southern blots of *Bam*HI-digested DNA from PC7183, we found that the aberrant κ rearrangement hybridized to the C_κ probe, but not to the J_κ probe, nor to a probe for the 5' portion of the intervening sequence (IVS). Subsequent mapping demonstrated that the rearrangement had occurred 3' of the *Hind*III site normally found in the intervening sequence. This is a different region from that observed in the plasmacytoma MPC-11³, in which rearrangement of a V seg-

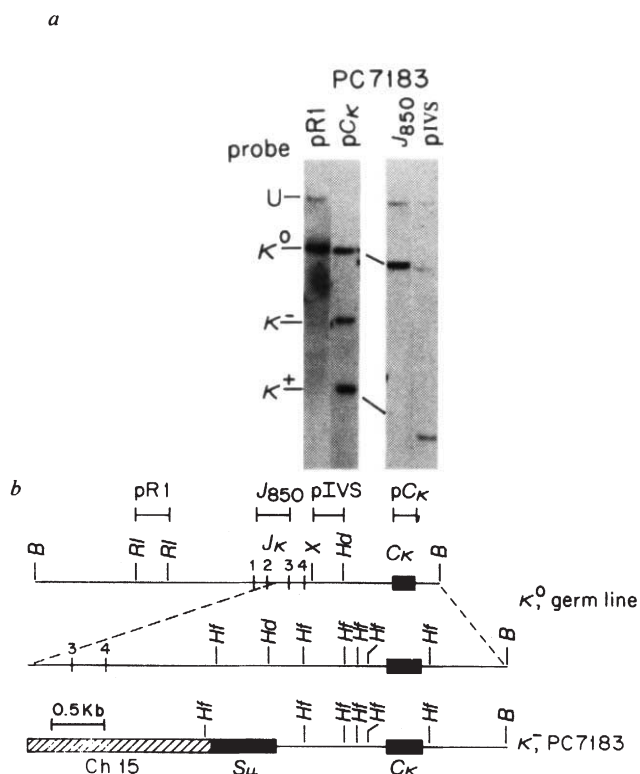


Fig. 1 *a*, Southern blot of *Bam*HI-digested DNA from PC7183 annealed with κ -specific probes (indicated by horizontal bars in panel *b*), showing the productively rearranged κ fragments (κ^+), the nonproductively rearranged κ fragment (κ^-) and a fragment containing sequences upstream of J_κ (U). The κ^+ rearrangement resulted from a join to J_4 , and thus does not hybridize to the J_{850} probe. The DNA in the last two lanes was run a greater distance on a second gel, thus accounting for the apparent size discrepancies. The band below the upstream band (labelled κ^0) is at the position of the germ-line κ fragment. The intensity of this band in this PC7183 DNA preparation is significantly higher than previously observed⁹ using PC7183 DNA prepared from the tumour at an earlier *in vivo* passage generation. This preparation may include more normal tissue contamination, or the tumour may have acquired a κ^0 gene during the course of propagation. The contexts of the κ^+ , κ^- and U fragments have remained the same in PC7183 preparations. *b*, The rearrangement of the κ^- allele, shown schematically. This rearrangement is the result of a recombination 3' of the κ *Hind*III site normally found in the IVS. This recombination involves chromosome 15 and a region homologous to the μ switch region (see Figs 2, 3). B, *Bam*HI; RI, *Eco*RI; X, *Xba*I; Hd, *Hind*III; Hf, *Hinf*I.

ment occurred within the IVS at a site having homology with J flanking recognition sequences. Because there are no other discernible joining sequences within the IVS we suggested that such a κ^- rearrangement may arise by a different kind of recombination event from that used in V-J joining.

The κ^- rearrangement in PC7183 was cloned as a 10 kilobase (kb) *Hind*III fragment into a λ KT1059-30 vector (gift of Dr K. Marcu). Using this 10 kb insert as a probe in Southern blots of *Eco*RI-digested germ-line DNA from several inbred mouse strains, three intense bands and one faint band were detected (Fig. 2a). One intense band (16 kb, labelled C_κ) can be attributed to the germ-line κ *Eco*RI fragment. Another intense band (8 kb) was common to all three strains, and the third varied in size among various inbred strains (14 kb in C57BL/6, 12 kb in BALB/c and DBA/2, 23 kb in A/HeJ, and 9.5 kb in C3H). The size distribution of the polymorphic fragment among 11 recombinant-inbred strains derived from BALB/c $\times$ C57BL/6 and C57BL/6 $\times$ DBA/2 matings agreed completely with the strain distribution pattern of genes on chromosome 12 (data not shown). We also noted that the sizes of the

§ Present address: Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242, USA.

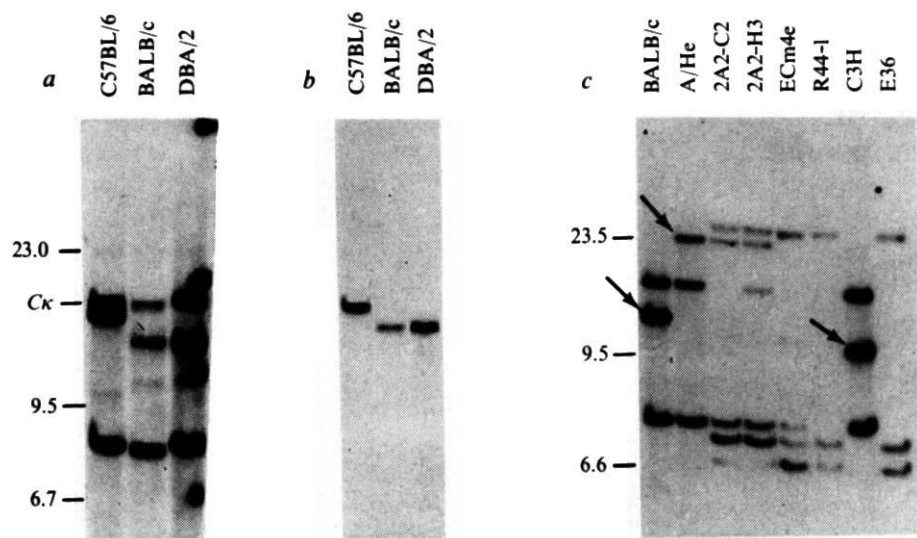


Fig. 2 Southern blot of *Eco*RI-digested DNA from various mouse strains and mouse/hamster hybrid cells probed with the cloned 7183 κ^- DNA (a, c) or pJ₁₄ (b). In c, E36 is the hamster parent of the hybrid cells; A/He is the mouse parent of 2A2-C2 and 2A2-H3; C3H is the mouse parent of ECm4e and R44-1 (see Table 1 for chromosomal assignments). Arrows point to the S_μ fragment, which varies in size among the three strains shown.

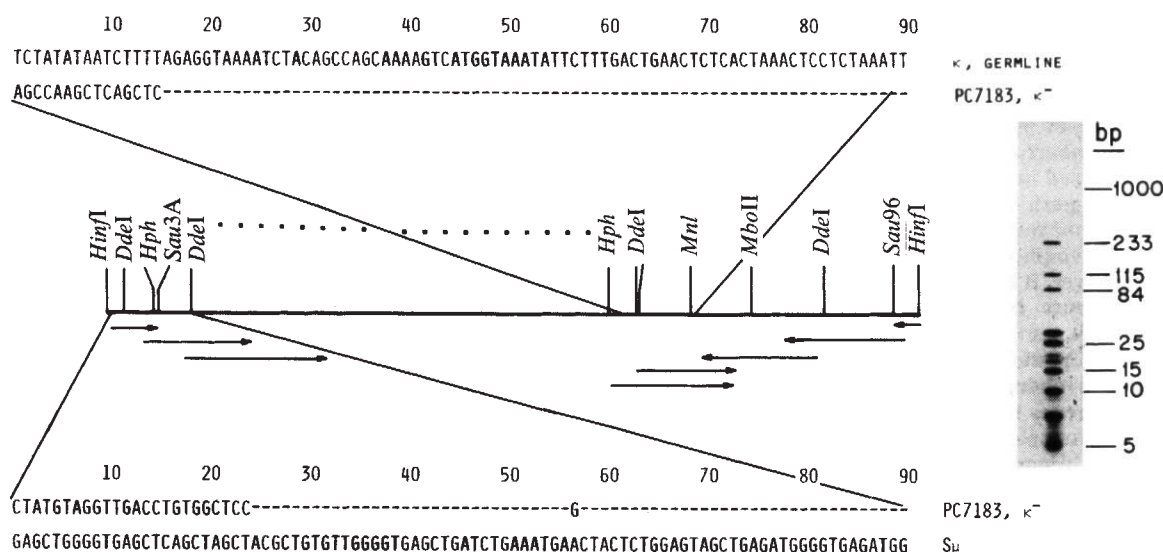


Fig. 3 Sequence comparison of the 1 kb *Hinf*I fragment from the κ^- of PC7183 with part of κ germ line¹³, and part of the μ switch region¹². Arrows show the extent and strategy of sequencing. All fragments were end-labelled with [γ -³²P]ATP and polynucleotide kinase, cut with a second restriction enzyme, and sequenced by the method of Maxam and Gilbert¹⁹. Only a portion of each sequence is shown to demonstrate regions of recombination. Sequence homology of the PC7183 κ^- with S_μ (lower) or κ (upper) is indicated with a dashed line; the sequences are shown where they differ. From the point of recombination the remainder of the upper sequence differed at only 2 out of 184 sites with the S_μ sequence. The sequence extending 3' of the recombination region shown on the right is identical with the κ germ-line sequence (346 bp). The region between *Dde*I sites (indicated with dots) contained multiple *Dde*I sites, typical of S_μ . This was demonstrated by digesting the *Hinf*I fragment with *Dde*I, end-labelling the fragments with [γ -³²P]ATP and polynucleotide kinase, and running the labelled DNA on a 10% polyacrylamide gel. The inset shows that, in addition to the *Dde*I fragments derived from the 5' and 3' ends (84, 115, 233 bp), there are multiple *Dde*I sites within the fragment which generate fragments 5–30 bp in length.

polymorphic band corresponded to those observed when *Eco*RI digests of various mouse strains were probed for the region of the heavy-chain gene that includes the μ switch (S_μ) region¹¹. To verify this apparent identity, the PC7183 DNA probe was washed off the blot in Fig. 2a, and the blot was rehybridized with pJ₁₄, a probe containing the first two coding domains of C_μ as well as part of the S_μ sequence¹¹. As shown in Fig. 2b the same polymorphic band was indeed observed. The faint band seen in Fig. 2a corresponds in size to an *Eco*RI fragment containing the α switch region, known to cross-hybridize with S_μ ¹¹. This band is also polymorphic among the various mouse strains.

To determine the chromosomal origin of the 8 kb fragment that was common to the three mouse strains tested, blots of DNA from mouse/hamster hybrid cell lines that contain different complements of mouse chromosomes were probed with the PC7183 cloned DNA. The results of a preliminary

blot of several mouse/hamster lines indicated that the 8 kb fragment originated from chromosomes 2, 15 or 17 (data not shown). To distinguish between these chromosomes, four mouse/hamster lines expressing different combinations of these chromosomes were probed (see Fig. 2c and Table 1). All three chromosomes are present in line 2A2C2; line 2A2-H3 lacks chromosome 2; line ECm4e lacks chromosomes 2 and 17; and line R44-1 lacks chromosomes 2 and 15. As shown in Fig. 2c, only line R44-1 does not contain the 8 kb fragment. As only chromosome 15 is common to all of the remaining lines, the assignment of the 8 kb fragment can be made to chromosome 15. The loss of C_κ or C_μ hybridization in some of the lines is expected as they lack chromosome 6 and 12 (see Table 1).

To determine the recombination site within the κ^- allele of PC7183, the clone was mapped and partially sequenced. By mapping the PC7183 clone in parallel with a κ germ-line clone, a 1 kb *Hinf*I fragment was found which spanned the κ recombination site.

Table 1 Chromosomal assignments of mouse/hamster somatic cell hybrids

Chromosome	Hybrid cell line			
	Mach 2A2-C2*	Mach 2A2-H3*	EC _{m4e} †	R44-1‡
2	0.60	0.00	0.00	0.00
6	0.00	0.80	0.00	0.00
12	0.90	1.40	0.00	0.00
14	0.00	1.15	1.00	0.00
15	0.35	1.30	1.00	0.00
17	0.85	1.05	0.00	1.00
8 kb‡ fragment	+	+	+	-

* Mouse parent A/HeJ; † mouse parent C3H/HeJ; ‡ from Fig. 2c.

nation site; the 3' *HinfI* site is derived from the κ locus and the 5' *HinfI* site from the foreign DNA associated with the κ locus (see Fig. 1). The *HinfI* fragment was then isolated for DNA sequencing. A total of 273 base pairs (bp) was sequenced from the 5' end, and 360 bp from the 3' end. Figure 3 shows a portion of the κ^- sequence, along with a portion of the S_μ sequence¹², and a portion of the κ germ-line sequence¹⁴. From the 5' end of the *HinfI* fragment, 24 of 40 bases of novel sequence is shown, immediately followed by a sequence (66 bases shown) which only differs at two sites from that published for S_μ (ref. 12). (The two differences might be due to polymorphisms between the strain in which PC7183 was induced, NZB, and the strain from which S_μ sequence was determined, BALB/c.) A 600 bp region within the 1 kb fragment was not sequenced for lack of adequate restriction sites. However, complete digestion of the 1 kb *HinfI* fragment with *DdeI* revealed the presence of tandem repetitive sequences typical of the S_μ region¹². The *HinfI* fragment was digested with *DdeI*, end-labelled with [γ -³²P]ATP and polynucleotide kinase, and run on a 10% polyacrylamide gel. As shown in Fig. 3 (inset), major bands were detected at 5, 10, 15, 20, 25 and 30 bp, consistent with our sequence data showing that the CTNAG *DdeI* recognition sequence is repeated at these intervals in the S_μ region. Also, fragments derived from the 5' end (84 bp) and 3' end (115 and 233 bp) were detected. No other fragments larger than 30 bp were detected, indicating that the core region of this fragment is comprised of the tandem repeats typical of S_μ . Thus, in agreement with the Southern blot analysis, the sequence and end-labelling data indicate that this region is derived from the S_μ region of chromosome 12. However, we cannot rigorously exclude the possibility that it is derived from a separate region of DNA highly homologous to S_μ . A recombination site at the position of nucleotide 24 in Fig. 3 (lower) would then presumably involve a join of DNA from chromosome 15 to the S_μ region. When a 1.7 kb *BamHI* fragment, derived from the 5' end of the 10 kb cloned insert, was used as a probe against a Southern blot for *EcoRI*-digested germ-line DNA, only the 8 kb fragment assigned to chromosome 15 was observed (not shown). Therefore, the entire 7 kb, from the 5' end of the clone to the S_μ region, is derived from chromosome 15.

At the 3' end of the *HinfI* fragment, 346 nucleotides were found to be identical with the κ germ-line sequence¹³. However, the first 15 nucleotides shown in Fig. 3 (upper) diverge from the κ sequence. Although there is no available S_μ sequence that can be compared with this region (its position places it within an unsequenced portion of the S_μ region)¹², the DNA does contain a *Hph* site and two *DdeI* sites within 25 bp, a characteristic of S_μ . We conclude, then, that the S_μ sequence extends through the unsequenced portion of the *HinfI* fragment to the point of recombination with the κ locus.

Although karyotypic evidence for translocation is not yet available, the PC7183 κ^- could be regarded as a double translocation, in which case two successive events must be considered.

An interesting alternative might involve transposition of the S_μ region to chromosome 15 followed by a translocation to the C_κ locus on chromosome 6. The sequence at the S_μ - C_κ junction does not reveal exactly how the recombination was generated; however, it may be significant that the recombination occurred within an A+T-rich region which is highly conserved between man and mouse¹⁴. What is particularly interesting is that translocations involving chromosome 15 and immunoglobulin bearing chromosomes have frequently been observed in tumour cells^{15,16}, and DNA from chromosome 15 has recently been observed adjacent to the α heavy-chain switch region of a plasmacytoma¹⁷. It is noteworthy that the chromosome 15 DNA of the PC7183 κ^- does not cross-hybridize with the chromosome 15 DNA observed in the α switch region (B.G.V.N. and K. Marcu, data not shown). These results raise the possibility that various regions on chromosome 15 may be involved in chromosomal translocations to immunoglobulin genes.

This form of recombination involving the κ locus may help to clarify certain unusual features observed in plasmacytomas. We previously demonstrated that the DNA between rearranging gene segments of the κ locus is frequently retained in plasmacytomas as well as normal B cells⁴. The DNA sequence of several retained DNA segments showed they are likely to be the reciprocal product of a V-J join (ref. 18 and B.G.V.N., unpublished). Because the reciprocal fragments examined are not related to the V and J joined in the cell in which they are found, we suggested that the mechanism of V-J joining may involve unequal sister chromatid exchange (USCE), whereby the V-J fusion segregates away from the reciprocal product. We found, however, an unusually high frequency of reciprocal rearrangement fragments in plasmacytomas with κ^+/κ^- genotypes, as compared with those with κ^+/κ^0 genotypes, and suggested that certain κ^- forms might arise by other recombinational mechanisms that do not necessarily result in segregation of products as expected in USCE. The translocational recombination leading to the κ^- rearrangement of PC7183 would be an example of such a mechanism. We have also observed some reciprocal fragments which do not appear to be of the type generated from USCE in V-J joining. Figure 1 shows such an example, in which a reciprocal fragment (U, in Fig. 1) includes IVS DNA. This might represent a non-segregating reciprocal of the κ^- rearrangement in PC7183. In addition, if tumour specific translocations contribute to the frequency of aberrant rearrangements of immunoglobulin loci, then one would expect to find a higher frequency of κ^+/κ^- genotypes in plasmacytomas, as compared with normal lymphocytes, as has actually been observed^{5,10}.

We thank Joselina Gatmaiten for technical assistance. This work was supported by grants from NIH and an appropriation from the Commonwealth of Pennsylvania. B.G.V.N. was supported by a fellowship from the Damon Runyon-Walter Winchell Cancer Fund.

Received 25 October; accepted 3 December 1982.

1. Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa, S. *Nature* **290**, 562-565 (1981).
2. Max, E. E. *et al. Cell* **21**, 793-798 (1980).
3. Seidman, J. G. & Leder, P. *Nature* **286**, 779 (1980).
4. VanNess, B. G., Coleclough, C., Perry, R. P. & Weigert, M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 262-266 (1982).
5. Coleclough, C., Perry, R. P., Karjalainen, K. & Weigert, M. *Nature* **290**, 372-378 (1981).
6. Kirsch, I. R. *et al. Nature* **293**, 585-587 (1981).
7. Harris, L. J., Lang, R. B. & Marcu, K. B. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4175-4179 (1982).
8. Leder, P. *et al. Cold Spring Harb. Symp. quant. Biol.* **45**, 859-865 (1980).
9. Perry, R. P. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
10. Perry, R. P., Coleclough, C. & Weigert, M. *Cold Spring Harb. Symp. quant. Biol.* **45**, 925-933 (1980).
11. Marcu, K. B. *et al. Cold Spring Harb. Symp. quant. Biol.* **45**, 899-911 (1980).
12. Nikaïdo, T., Nakai, S. & Honjo, T. *Nature* **292**, 845-848 (1981).
13. Max, E., Maizel, J. & Leder, P. *J. biol. Chem.* **256**, 5116-5120 (1981).
14. Heiter, P. A. *et al. Cell* **22**, 197-207 (1980).
15. Ohno, S. *et al. Cell* **18**, 1001-1007 (1979).
16. Klein, G. *Nature* **294**, 313-318 (1981).
17. Harris, L. J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
18. Höcht, J., Müller, C. R., Zachau, H. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1383-1387 (1982).
19. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1979).

Igh-V or closely linked gene(s) control immunological memory to a thymus-independent antigen

J.-H. Colle, I. Motta, B. Shidani & P. Truffa-Bachi

Unité d'Immunophysiologie Moléculaire, ER 232 CNRS, Institut Pasteur, 75724 Paris Cédex 15, France

Mice mount a normal primary antibody response on stimulation with the thymic-independent antigen trinitrophenylated lipopolysaccharide (TNP-LPS). Although we have previously reported the generation of functional B-memory lymphocytes to TNP-LPS^{1,2}, this memory response was only observed in few mouse strains³. Here we have used congenic mouse strains in an attempt to locate the genetic regions involved in the memory response. We show that genes of the major histocompatibility complex (MHC) do not have a critical role but that genes coding for the variable region of immunoglobulin heavy chains or gene(s) closely linked to them are required for memory cell induction by TNP-LPS.

The capacity of an animal to mount an immune response to a given antigen depends on genetic control inherited in a mendelian fashion^{4,5}. In the mouse, the genes controlling immune responses were shown to belong to the major histocompatibility *H-2* complex and were named *Ir* genes^{6,7}. Inbred mouse strains could be separated into responders and nonresponders to a particular antigen⁸. Nonresponder mice produce a normal IgM primary response and little, if any, IgG antibodies or memory cells; responder mice develop an immunological memory and a strong IgG response⁶. The control of specific immune responses by *Ir* genes is restricted to thymus-dependent antigens and in no instance has control of the response to thymic-independent antigens by *Ir* genes been recorded⁹.

However, we have recently demonstrated that an immunological memory response, similar in many features to the one obtained with thymus-dependent antigens, is induced by TNP-LPS in some mouse strains but not in many others⁷⁻⁹. That is, all strains produce an IgM primary response, but only some (for example C57BL/6) develop a TNP-specific immunological memory to TNP-LPS, reflected in a high IgG response observed after a challenge^{1,2}.

To investigate whether or not genes of the *H-2* complex are implicated in this genetic control, congenic mouse strains possessing the genetic background of the responder strain (C57BL) and the *H-2* haplotype derived from nonresponder strains (DBA/2, AKR) were tested for their ability to develop a memory response to TNP-LPS. Conversely, we have studied the effect of the *H-2^b* haplotype, transposed from a responder strain (C57BL/6) into a nonresponder strain (BALB/c). The kinetics of primary and secondary anti-TNP responses for the C57BL/10, DBA/2, AKR, BALB/c and the congenic B10.D2, B10.BR and BALB.B mouse strains are shown in Figs 1 and 2. Whereas DBA/2, AKR, BALB/c and BALB.B mice display a secondary response which can be even lower than the primary response, C57BL/10, B10.D2 and B10.BR mice give rise to a secondary anti-TNP response which is four to five times higher than the primary response and characterized by a large number of cells secreting antibodies of the IgG isotype. Thus, the presence of a (nonresponder) *H-2^d* or *H-2^k* haplotype on a C57BL genetic background does not affect the development of memory response to TNP-LPS. In addition, BALB.B mice which possess the (responder) *H-2^b* haplotype from C57BL/6 strain fail to develop an immunological memory to TNP-LPS.

As we were unable to find a direct influence of the *Ir* genes on the secondary response to TNP-LPS, we investigated the genes coding for the heavy chains of immunoglobulins¹⁰⁻¹²

which are involved in the regulation of the immune response. Light chain loci are also involved and these genes are known to display a certain degree of polymorphism¹²⁻¹⁴ but they have not been investigated here. The congenic strains CB20 and BAB14 possess the same genetic background as BALB/c, a nonresponder strain, but differ from it in a fragment of DNA on chromosome 12; in CB20, the genes coding for the constant *Igh-C* and variable *Igh-V* regions of the immunoglobulin heavy chains originate from C57BL/6 (ref. 12); in BAB14, the *Igh-C* locus and only a part of the *Igh-V* genes derived from C57BL/6 (ref. 12). The comparison of the kinetics of the secondary anti-TNP response to TNP-LPS obtained in CB20 and BAB14 mice (Fig. 2) indicates that the *Igh-C* genes or genes located on the prealbumin side of the chromosome are not associated with the regulation of the secondary TNP-LPS response because BAB14, like BALB/c mice, are nonresponders. In contrast, CB20 mice display an unequivocal immunological memory to TNP-LPS, indicating that the *Igh-V* or closely linked genes, located between *Igh-V* genes and the centromere, are implicated in the ability to develop a memory response to thymus-independent antigens.

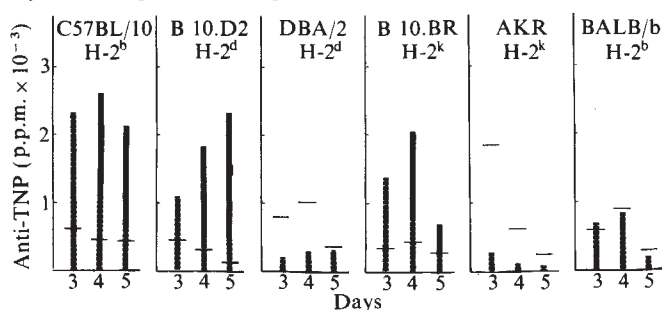


Fig. 1 Effect of different *H-2* haplotypes on the secondary anti-TNP-LPS immune response. B10.D2 mice were obtained from Olac, and B10.BR and BALB.B were obtained from CNRS. All the other strains were bred in our own department. Male or female mice, 2–3 months old, were selected for investigation. Mice were primed intraperitoneally with 5 µg of TNP-LPS and boosted with the same dose of antigen 7 days later. The anti-TNP response was determined by counting the cells secreting anti-TNP antibodies in the spleen, 3, 4 and 5 days after the challenge. Plaque-forming cells (PFC) were counted by the method developed by Cunningham and Szenberg²¹, using lightly haptenated horse erythrocytes²² and guinea pig serum as source of complement. Indirect PFC were revealed with a rabbit anti-mouse immunoglobulin antiserum (a gift from Dr Bordenave) at 1:1,000 dilution. IgG PFC are expressed as the difference between PFC developed with complement+antiserum and PFC detected with complement alone. Results are expressed in PFC per 10⁶ nucleated cells (counted in a model D Coulter counter). Each bar represents the mean of 3–15 mice per group. The variability did not exceed 20% of the mean. Secondary anti-TNP response: ▨▨▨, IgM; ▨, IgG. The horizontal bars represent the primary anti-TNP response.

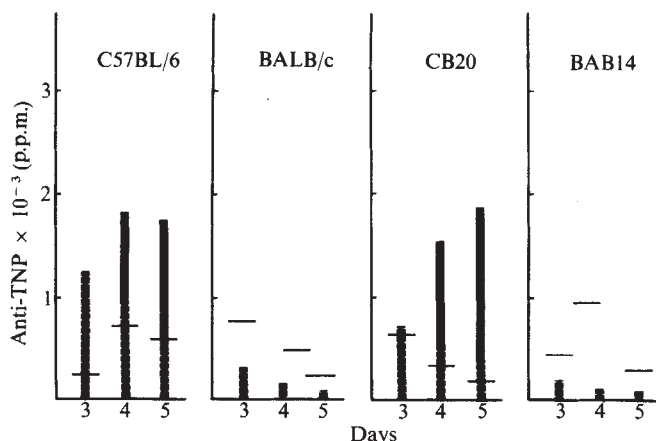


Fig. 2 Effect of genes located on chromosome 12 on the secondary anti-TNP-LPS immune response. Methods are as described in Fig. 1 legend.

If *Igh-V* genes, rather than the closely linked genes, are necessary for the memory response to TNP-LPS, then idiotype-anti-idiotype interactions¹⁵ may be involved in the induction and/or the appearance of IgG memory cells, a phenomenon which has not previously been described. The fact that memory response can be observed even in T-cell-depleted animals² is not an argument against this hypothesis; different authors^{16,17} have reported that B-B-cell interactions can successfully modulate the expression of particular cell clones. Alternatively, as the B-memory cells are clearly distinct from primary B cells, they may belong to an independent B-cell lineage^{18,19}. The existence of genes governing the ontogeny of B cells is well documented (reviewed in ref. 20); however, nothing is known about genes controlling the emergence of B-memory cells. If genes distinct from the *Igh-V* genes govern the expression of the memory response to TNP-LPS, their key role may be in the development of B-memory cells. Experiments are being performed to distinguish between these two hypotheses.

We thank Dr G. Bordenave for helpful discussions, Drs P.A. Cazenave and M. Seman for the CB20 and BAB14 mice and Mr R. Perret for technical assistance. This work was supported by grants from the Institut Pasteur and CNRS (ER232 and ATP 50-50).

Received 3 September; accepted 29 November 1982.

1. Motta, I. & Truffa-Bachi, P. *Ann. Immun., Inst. Pasteur* **131C**, 223-227 (1980).
2. Motta, I., Portnoi, D. & Truffa-Bachi, P. *Cell Immun.* **57**, 327-338 (1981).
3. Colle, J. H., Motta, I. & Truffa-Bachi, P. in *Mechanisms of Lymphocytes Activation* (eds Resch, K. & Kirchner, H. 491-493 (Elsevier, Amsterdam, 1981).
4. Kantor, F. S., Ojeda, A. & Benaceraff, B. *J. exp. Med.* **117**, 55-69 (1963).
5. Levine, B. B. & Benaceraff, B. *Science* **147**, 517-518 (1965).
6. McDevitt, H. O. & Sela, M. *J. exp. Med.* **122**, 517-531 (1965).
7. Benaceraff, B. & McDevitt, H. O. *Science* **175**, 273-279 (1972).
8. Benaceraff, B. *Ann. Immun., Inst. Pasteur* **125C**, 143-164 (1974).
9. Benaceraff, B. & Germain, R. N. *Immun. Rev.* **38**, 101-119 (1978).
10. Blomberg, B., Geckeler, W. R. & Weigert, M. *Science* **177**, 178-180 (1972).
11. Eichmann, E. *Eur. J. Immun.* **4**, 301-307 (1972).
12. Lieberman, R. in *Springer Semin. Immunopath.* **1**, 7-30 (1978).
13. Gottlieb, P. D. *J. exp. Med.* **140**, 1432-1437 (1974).
14. Gibson, D. *J. exp. Med.* **144**, 298-303 (1976).
15. Jerne, N. K. *Ann. Immun., Inst. Pasteur* **125C**, 373-389 (1974).
16. Nisonoff, A. in *Cells of Immunoglobulin Synthesis* (ed. Pernis, B.) 335-342 (Academic, New York, 1979).
17. Primi, D., Juy, D. & Cazenave, P. A. *Eur. J. Immun.* **11**, 393-398 (1981).
18. Klinman, N. R. & Press, J. L. *Transplant Rev.* **24**, 41-83 (1975).
19. Teale, J. M., Lafrenz, D., Klinman, N. R. & Strober, S. *J. Immun.* **126**, 1952-1957 (1981).
20. Klinman, N. R., Wylie, D. E. & Teale, J. M. *Immun. Today*, November, 212-219 (1981).
21. Cunningham, A. J. & Szenberg, A. *Immunology* **14**, 599-601 (1968).
22. Rittenberg, M. B. & Pratt, K. L. *Proc. Soc. exp. Biol. Med.* **132**, 575-581 (1969).

Identity of some human bladder cancer cell lines

Carol M. O'Toole*, Susan Povey†, P. Hepburn‡
& L. M. Franks§

* Division of Immunology, Department of Pathology, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QQ, UK
† MRC Human Biochemical Genetics Unit, The Galton Laboratory, University College, London NW1 2HE, UK
‡ Department of Pathology, Institute of Urology, St Paul's Hospital, London WC2H 9AE, UK
§ ICRF, Lincoln's Inn Fields, London WC2A 3PX, UK

Recent reports on transfection of mouse cells with DNA from the established human urinary bladder cancer cell lines T24, J82 and EJ (MGH-U1)¹⁻⁴, and the presence of an identical genetic modification in T24 and EJ cells^{5,6} have led us to examine the identity of these and other cultures of urothelial origin. By the criteria of HLA-A-B-C typing⁷ and isozyme analysis⁸, we conclude that EJ (MGH-U1) and some cultures of J82 are in fact T24 cells. However, five other bladder cancer cell lines, J82 (CO'T), RT4, RT112, TCCSuP and SCaBER, are clearly distinct from T24 by HLA typing (ref. 7) and/or isozyme⁸ patterns.

The HLA phenotypes of the cell lines we have studied are given in Table 1. The HLA profile of the carcinoma line J82(CO'T), originated and maintained by CO'T, showed full agreement with that of the donor's peripheral blood lymphocytes (PBL) and of an Epstein-Barr virus-transformed cell line, J82 EBV, from the same individual.

The HLA profiles of the six other lines including T24 and EJ (hereafter referred to as MGH-U1) were identical. Lymphocytes from the donor of one of the lines, Hu961T, have been typed as HLA-A2, A25/B7, B18 (ref. 9). The donors of the other four lines were not HLA-typed.

Isozyme data on all lines are summarized in Table 2. The lines which had been cultured for a long time at the Imperial Cancer Research Fund (indicated as LMF) were tested both in recent passage and from the earliest frozen stocks. Many clones of MGH-U1 have also been tested.

The results confirm that the J82 (CO'T) line is derived from the same individual as J82 EBV, the best direct evidence being the presence of the acid phosphatase (ACP1) CA phenotype, seen in only about 4% of Europeans⁸. The apparent homozygosity of J82 (CO'T) for two loci on chromosome 6, which are heterozygous in expression in J82 (EBV), is not particularly surprising as increased homozygosity on prolonged culture is found in some other tumours, especially lymphomas¹⁰. J82 (CO'T) was established *in vitro* in 1972 (ref. 11), while J82 EBV has been cultured since 1980.

All the other lines are strikingly similar for all enzymes tested and at all passages and in all clones (data not shown in detail), except that the ACP1 phenotype of MGH-U1 and J82(LMF) is typed as BA whereas they were originally recorded, perhaps in error, as type B (ref. 8). We would now type all these lines as BA for ACP1, although in some clones of MGH-U1 the B component is very much stronger than the A. Although no individual enzyme phenotype is uncommon the combination seen in T24 would be expected in less than 1% of the population.

We also have some data on the tumorigenicity and karyotype of the cell lines. T24 (LMF) cells had a chromosome mode of 84 and MGH-U1 cells had a mixed population with subclones having modes of 46 and 90/92 (ref. 12). There were also major differences in tumorigenicity in nude mice (4-10 × 10⁶ cells, observed for 6 months). T24 (LMF) cells produced tumours in 8 of 51 mice. In a separate experiment, using T24 (CO'T) cells from a laboratory which has never had MGH-U1 cells, no tumours were produced in five mice. J82 (LMF) cells gave one tumour in 14 mice. MGH-U1 cells produced 44 tumours in 54 mice, within 1-2 months, although some clones of MGH-U1 were not tumorigenic. One-dimensional polyacrylamide gel analysis of whole cell proteins have shown no differences between T24 (LMF) and MGH-U1 parent cell lines, although there are minor differences in the clones and in some of the cell surface proteins (unpublished data).

DNA from T24 cells produces high levels of transfection in the mouse line NIH 3T3 (refs 1-4). Some authors have also reported efficient transfection with DNA from J82 (ref. 1) and MGH-U1 (refs 1, 3) cells. However, others have failed to be transfected with J82 DNA (ref. 4). Our data suggest that the explanation for these discrepancies lies in the cross-contamination of some cultures of J82 and MGH-U1 with T24 cells. The data reported by Der *et al.*¹ were obtained with cultures of 'J82' and MGH-U1 derived from L.M.F.'s laboratory (G. Cooper, personal communication). However, the data reported by Perucho *et al.*⁴ in which J82 DNA failed to transfect⁴ were obtained from J82 cells supplied to Sloan Kettering Institute by C.O'T. The presence of a common genetic lesion in T24 and EJ (MGH-U1)^{5,6} together with our data suggests that these lines are, in fact, identical.

The combined HLA and isozyme results of all the carcinoma lines reported here, except for J82 (CO'T), make it highly probable that they are derived from the same individual. The possibility that J82 (LMF) and MGH-U1 are identical has been considered previously⁸; at that time T24 was thought to differ from those at one locus. However the present results suggest

Table 1 Origin and HLA phenotype of bladder cancer cell lines

Cell Line	Refs	HLA on donor PBL*	HLA on cell line†	Source of culture
T24 (CO'T)‡	13	Unknown	A1, A3/B18/Cw5	C.O.T.
T24 (LMF)	13	Unknown	A1, A3/B18/Cw5	L.M.F.
J82 (CO'T)	11	A2, Aw32/Bw51(5), Bw44(12)/Cw5	A2, Aw32/B5, B12/Cw5	C.O.T.
J82EBV		A2, Aw32/Bw51(5), Bw44(12)/Cw5	A2, Aw32/B5, B12/Cw5	C.O.T.
J82(LMF)		A2, Aw32/Bw51(5), Bw44(12)/Cw5	A1, A3/B18/Cw5	L.M.F.
EJ (MGH-U1)	17	Unknown	A1, A3/B18/Cw5	L.M.F.
MGH-U2	17	Unknown	A1, A3/B18/Cw5	P.H.
Hu456	9	Unknown	A1, A3/B18/Cw5	P.H.
Hu961T	9	A2, A25/B7, B18	A1, A3/B18/Cw5	P.H.

* Peripheral blood lymphocytes.

† Cells were HLA-typed by antibody-dependent cell-mediated cytotoxicity as described in ref. 7.

‡ Obtained from originator, Dr J. Bubenik¹³.**Table 2** Enzyme data on various bladder cancer cell lines

Cell line	PGM1	PGM3	GOT2	ESD	AK1	ADA	ACP1	PGD	G6PD	PEPA	PEPC	GLO	FUC	PGP	GDH
T24 (CO'T)	1	1	1	1	1	1	BA	A	B	—	—	1	2-1	1	1
T24 (LMF)	1	1	1	1	1	1	BA	A	B	1	1	1	2-1	1	1
J82 (CO'T)	1	2	1	1	1	1	CA	A	B	—	—	2	2-1	1	1
J82 EBV	1	2-1	1	1	*	1	CA	A	B	—	—	2-1	2-1	1	1
J82 (LMF)	1	1	1	1	1	1	BA†	A	B	1	1	1	2-1	1	—
EJ (MGH-U1)	1	1	1	1	1	1	BA†	A	B	1	1	1	2-1	1	1
MGH-U2	1	1	1	1	—	1	BA	A	B	—	—	1	2-1	1	1
Hu 456	1	1	1	1	—	1	BA	A	B	—	—	1	2-1	1	1
Hu 961T	1	1	1	1	—	1	BA	A	B	—	—	1	2-1	1	1
Frequency of phenotype of T24 in European population	0.57	0.55	0.97	0.82	0.92	0.9	0.42	0.96	1.0	1.0	0.95	0.44	0.4	0.7	0.49

Enzymes typed by horizontal starch gel electrophoresis as in ref. 14 except for α -fucosidase (FUC) and glucose dehydrogenase (GDH) which were examined by isoelectric focusing as in refs 15 and 16, respectively. Other abbreviations: PGM1, PGM3, first and third loci of phosphoglucomutase; GOT2, mitochondrial glutamate-oxaloacetate transaminase; ESD, esterase D; AK1, adenylate kinase; ADA, adenosine deaminase; ACP1, first locus of acid phosphatase; PGD, 6-phosphogluconate dehydrogenase; G6PD, glucose-6-phosphate dehydrogenase; PEPA and PEPC, peptidases A and C; GLO, glyoxalase; PGP, phosphoglycolate phosphatase.

* Not expressed in lymphoid lines.

† Previously typed as 'B' (see text).

that it is T24 which has contaminated the others, as this line was established before the others and T24 (CO'T) was obtained directly from the originator. It would clearly be of interest to test an 'original' MGH-U1 and we hope to do so.

After many years in culture it is often difficult to establish the origin of tumour lines by comparison between sublines. Although our conclusion is that all the cells studied here are derived from two individuals, this does not mean that all sublines are identical. Major differences in tumorigenicity exist between MGH-U1 and T24 (LMF) (ref. 12) and also between the different clones of MGH-U1. The J82 (LMF) and T24 (LMF) cells have similar karyotypes, however J82 (CO'T) differs from both in karyotype particularly in the presence of two Y chromosomes¹¹. There is a large degree of chromosomal variation within MGH-U1 and between this and T24 (ref. 12), although the enzyme phenotypes have been stable. However, loss of function of alleles giving increased apparent homozygosity, such as has occurred in J82 (CO'T), is relatively common in tumours after prolonged *in vitro* culture, whereas alteration of mobility to give a 'new' isozyme is rare¹⁰. In the cell lines described here, the HLA phenotypes have remained stable although it would be helpful to know the HLA type of the original T24 donor. As we have suggested previously, the problems of cell line identification encountered here would be solved much more efficiently if normal tissue (lymphocytes and red cells) were obtained from the donor and stored frozen. In the event of a line being successfully established, this material could then be used for HLA and isozyme analysis and the results could usefully be included in the first published description of the line.

These results show that the cross-contamination of cell cultures by established lines is a persistent problem not restricted to HeLa cells, and that lines used for comparative studies such as the detection of oncogenes¹⁻⁶ or as targets for cell-mediated immunity^{9,11,13} should be frequently monitored for HLA and/or isozyme patterns.

We thank Drs G. Cooper, B. Christensen, G. Prout and M. Vilien for permission to publish these data, Drs Eva Wolf, Valerie Joysey and Julia Bodmer for gifts of HLA antisera, and M. Parkar for technical assistance.

Received 22 October; accepted 2 December 1982.

- Der, C. J., Krontiris, T. G. & Cooper, G. M. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3637-3640 (1982).
- Goldfarb, M., Shimizu, K., Perucho, M. & Wigler, M. *Nature* **296**, 404-409 (1982).
- Parada, L. F., Tabin, C. J., Shih, C. & Weinberg, R. A. *Nature* **297**, 474-478 (1982).
- Perucho, M. *et al. Cell* **27**, 467-476 (1981).
- Tabin, C. J. *et al. Nature* **300**, 143-149 (1982).
- Reddy, E. P., Reynolds, R. K., Santos, E. & Barbacid, M. *Nature* **300**, 149-152 (1982).
- O'Toole, C. M., Tiptaft, R. C. & Stevens, A. *Int. J. Cancer* **29**, 391-395 (1982).
- Povey, S., Hopkinson, D. A., Harris, H. & Franks, L. M. *Nature* **264**, 60-63 (1976).
- Troye, M., Vilien, M., Paper, G. R. & Perlmann, P. *Int. J. Cancer* **25**, 33-43 (1980).
- Povey, S., Jeremiah, S., Arthur, E., Steel, M. & Klein, G. *Ann. hum. Genet.* **44**, 119-133 (1980).
- O'Toole, C., Price, Z. H., Ohnuki, Y. & Unsgaard, B. *Br. J. Cancer* **38**, 64-76 (1978).
- Hastings, R. & Franks, L. M. *Int. J. Cancer* **27**, 15-21 (1981).
- Bubenik, J. *et al. Int. J. Cancer* **11**, 765-773 (1973).
- Harris, H. & Hopkinson, D. A. *Handbook of Enzyme Electrophoresis in Human Genetics* (Elsevier, Amsterdam, 1976).
- Turner, B. M., Turner, V. S., Beratis, N. G. & Hirschhorn, K. *Am. J. hum. Genet.* **27**, 651 (1975).
- King, J. & Cook, P. J. L. *Ann. hum. Genet.* **45**, 129-134 (1981).
- Kato, T., Irwin, R. I. & Prout, G. R. *J. exp. Med.* **157**-164 (1977).

Enkephalin reduces quantal content at the frog neuromuscular junction

J. L. Bixby & N. C. Spitzer

Department of Biology, University of California, San Diego, La Jolla, California 92093, USA

Opiate peptides, particularly Met-enkephalin, have been shown to block or reduce Ca^{2+} -dependent events in a variety of neurones¹⁻⁶. Our own observations on the effect of enkephalin on developing Rohon-Beard neurones in *Xenopus* spinal cord¹³ suggested that enkephalin might interact directly with Ca^{2+} channels in vertebrate neurones. This possibility prompted us to look for an effect of enkephalin on another population of neuronal Ca^{2+} channels—those in the presynaptic terminals at the frog neuromuscular junction. We were encouraged to try this preparation by an early report that morphine, in rather high concentrations, reduced the amount of acetylcholine (ACh) released from frog muscle by nerve stimulation⁷. Our present results indicate that Met-enkephalin reversibly and specifically reduces the quantal content of transmitter release from nerve terminals in the frog cutaneous pectoris muscle, probably by means of an effect on inward Ca^{2+} current.

Our initial experiments examined the effect of Met-enkephalin on endplate potential (e.p.p.) amplitudes in conditions of low quantal content⁸. Enkephalin (10–30 μM) applied by pressure ejection through a 'puffer' pipette led to a consistent but variable decrease in the size of the evoked response (Fig. 1a). This decrease averaged 40% ($n = 13$) and was not seen when normal saline was substituted for that containing enkephalin ($n = 3$), or when the enkephalin was applied in the presence of 15 μM naloxone ($n = 5$). Enkephalin at these concentrations often led to a depolarization of the muscle fibre membrane. This depolarization, when detectable, varied over 1–25 mV in amplitude and was not abolished by 15 μM naloxone. The effect of enkephalin on e.p.p. amplitude was independent of the presence or size of this depolarization.

These results are consistent with the hypothesis that enkephalin decreases Ca^{2+} entry into the presynaptic terminal, but there are several alternative explanations. To examine the possibility that the opiate was exerting its effect postsynaptically, we tested the effect of puffer-applied enkephalin on the ACh sensitivity of the endplate membrane. Enkephalin not only did not reduce, but in fact slightly increased the size of the response to iontophoretically applied ACh (Fig. 1b; $n = 3$). This increase in apparent sensitivity is in accord with the results of Haynes and Smith⁹, who found that opiate peptides inhibit 16S acetylcholinesterase (AChE) activity at the rat neuromuscular junction. When we repeated the iontophoretic application of ACh in the presence of neostigmine, we found no effect of enkephalin on the size of the ACh response in the two muscle fibres tested. This suggests that the reduction of e.p.p. size produced by enkephalin is due to a presynaptic action. As inhibition of AChE by enkephalin could complicate interpretation of our results, we did all subsequent experiments in saline containing neostigmine to block AChE activity.

It seemed possible that enkephalin was acting to decrease the size or duration of the presynaptic action potential, thus reducing Ca^{2+} entry indirectly, through an effect on the voltage-dependent Na^+ conductance. To examine this possibility, we performed experiments using focal stimulation of nerve terminals in the presence of tetrodotoxin (TTX)¹⁰. In this way, we could examine the effect of enkephalin on transmitter release produced by a controlled depolarization, as well as confine release to a relatively small region of the terminal, thus reducing contributions of distant regions which might not be affected by the enkephalin puff. Enkephalin applied to terminals stimulated

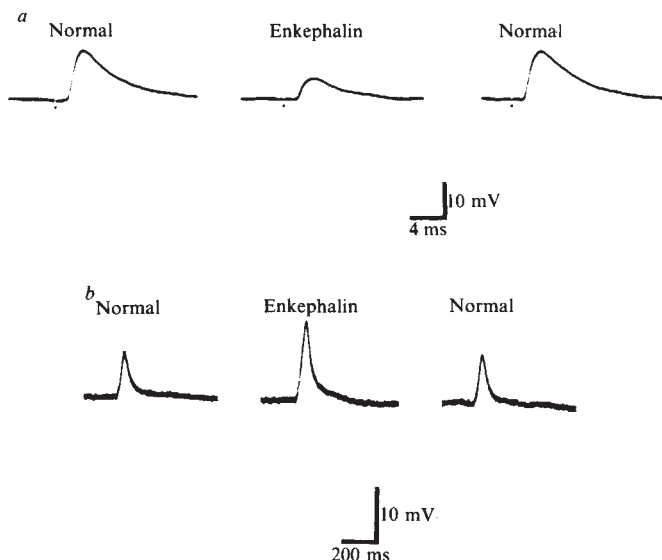


Fig. 1 *a*, Compound nerve stimulation with a suction electrode elicits an endplate potential that is reduced by focal puffer application of 20 μM Met-enkephalin. Saline contained 0.8 mM Ca^{2+} and 8.0 mM Mg^{2+} . *b*, Responses to iontophoretically applied ACh in the absence of neostigmine; the amplitude of the response is increased by focal application of 20 μM Met-enkephalin. Note that the response in enkephalin is longer as well as larger. This increase was abolished by the presence of 5×10^{-6} g ml^{-1} neostigmine (not shown).

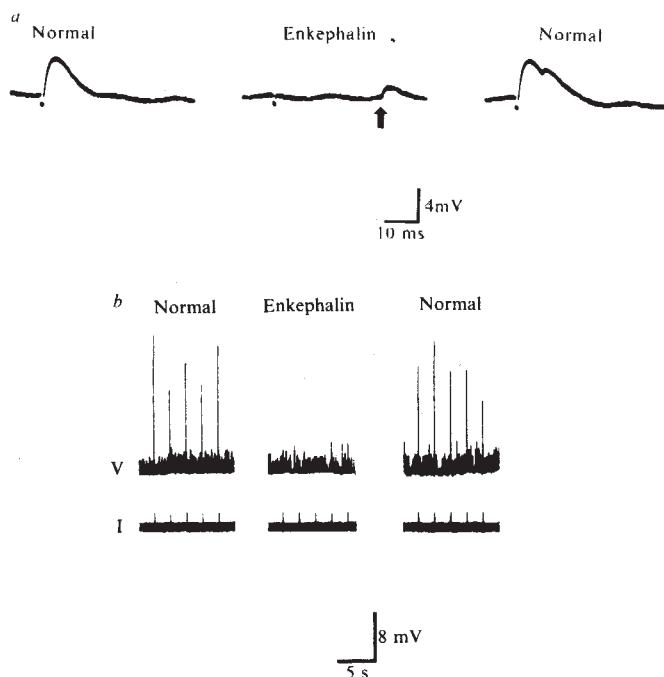


Fig. 2 *a*, Focal stimulation of a nerve terminal with an extracellular pipette elicits an endplate potential that is abolished by puffer application of 15 μM Met-enkephalin. Arrow indicates spontaneous miniature endplate potential (m.e.p.p.), apparently unaffected by application of the opiate. A second m.e.p.p. can be seen occurring during the right-hand e.p.p. *b*, Responses to repeated focal stimulation of a nerve terminal (left trace) are abolished by application of 15 μM Met-enkephalin by perfusion (1.5 min; centre trace) or from a puffer (not shown), while m.e.p.s continue unabated. There is no desensitization of the effect of Met-enkephalin, for periods up to 1 min. Response returns after 1 min in normal saline (right trace). Both bathing and puffer solutions contained neostigmine (5×10^{-6} g ml^{-1}). TTX (10^{-6} g ml^{-1}) was present in the experiment shown in *b*.

in this way, either by puffer application (10 or 15 μM) or by perfusion in the bath (15 or 20 μM), produced a dramatic decrease in the size of the postsynaptic response (Fig. 2b; mean 81%, $n = 11$). This was not due to a nonspecific effect of the TTX, as fibres examined using low currents to stimulate the terminals in the absence of TTX gave similar results (Fig. 2a). The advantage of delivery by puffer lies in the rapidity with which the local concentration can be changed; the advantage of bath application is that the perfusion pressure remains constant when solutions are exchanged. This reduction of e.p.p. size was specific, in that enkephalin applied in conjunction with naloxone (10 μM) had no consistent effect on the size of the focally stimulated e.p.p. ($n = 4$, puffer application; $n = 2$, bath application) when tested on nerve terminals shown to be affected by enkephalin alone.

The reduction in size of the evoked response could in principle be due to a decrease in either quantal size or quantal content, or both. This issue was addressed by examining the amplitude of spontaneous events in control conditions and during application of enkephalin. As can be seen from Fig. 2b, there was no obvious change in the amplitude of intracellularly recorded miniature endplate potentials (m.e.p.ps) during puffer application of enkephalin, but this could reflect contributions of quanta released from sites distant from the regions of opiate delivery. We therefore recorded miniature endplate currents (m.e.p.cs) focally, with an extracellular pipette¹¹, and examined the effect of puffer-applied enkephalin on the amplitude of these local events. The amplitude histograms indicate no change in the mean size of m.e.p.cs ($n = 3$). We obtained the same result examining intracellularly recorded m.e.p.p. amplitudes before and during bath application of enkephalin ($n = 3$); there also was no consistent change in the frequency of m.e.p.ps with bath-applied enkephalin. This indicates that enkephalin does not affect quantal size, and supports observations made with iontophoretically applied ACh in suggesting no change in the sensitivity of the postsynaptic membrane.

The results of these experiments indicate that Met-enkephalin reversibly and specifically reduces quantal content at the frog neuromuscular junction. The most likely mechanism involved is a block of voltage-dependent Ca^{2+} channels in the presynaptic terminal. We cannot rule out the possibility, however, that enkephalin reduces Ca^{2+} entry indirectly, by increasing a K^{+} conductance in the terminal. We do not know what concentrations of enkephalin are required to produce this effect because connective tissue in the muscle presents a diffusion barrier. Nevertheless, 15 μM is an upper limit to the concentration required for essentially complete blockade of transmitter release. It has been reported¹² that β -endorphin, a related opiate peptide, is present during development in rat motor neurones, where it has been suggested to play a part in the development of muscle AChE activity. If opiates are also present in frog muscle, they could be affecting either AChE activity, transmitter release, or both. It is tempting to postulate a physiological role for the effect of enkephalin on ACh release, perhaps mediated through an as yet undiscovered peptidergic innervation of muscle. However, the possibility also exists that enkephalin can be demonstrated to have an effect on neuronal Ca^{2+} channels in general, but that only some of these effects are of physiological relevance.

We thank David Van Essen and Alan Grinnell for comments on the manuscript, D. Van Essen for suggesting the focal stimulation experiments, Jorge Mancillas for his kind gift of naloxone, Rosario de Baca for technical assistance, and Martin Smith for discussions. This work was supported by grants from the NIH (NS15918) and the ONR (N00014-79-C-0748).

5. Illes, P., Zieglerberger, W. & Herz, A. *Naunyn-Schmiedeberg's Archs Pharmacol.* **308**, Suppl., R39 (1979).
6. Wouters, W. & Van der Bercken, J. *Neuropharmacology* **19**, 237-244 (1980).
7. Frederickson, R. C. A. & Pinsky, C. *Nature new Biol.* **231**, 93-94 (1971).
8. Grinnell, A. D. & Herrera, A. A. *J. Physiol. Lond.* **307**, 301-317 (1980).
9. Haynes, L. W. & Smith, M. E. *Neuroscience* **7**, 1007-1013 (1982).
10. Katz, B. & Miledi, R. *Proc. R. Soc. B* **167**, 23-38 (1967).
11. Fatt, P. & Katz, B. *J. Physiol., Lond.* **117**, 109-128 (1952).
12. Haynes, L. W., Smith, M. E. & Zakarian, S. *Brain Res.* **232**, 115-128 (1982).
13. Bixby, J. L. & Spitzer, N. C. *J. Neurosci.* (in the press).

Recruitment of cytosolic proteins to a secretory granule membrane depends on Ca^{2+} -calmodulin

M. J. Geisow & R. D. Burgoyne

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

An increase in free calcium triggers catecholamine secretion from chromaffin cells and calmodulin is strongly implicated as the intracellular Ca^{2+} receptor¹⁻³. In our recent studies of calmodulin action in the chromaffin cell, micromolar Ca^{2+} concentrations resulted in calmodulin^{4,5} and cytosolic proteins^{6,7} becoming bound to the chromaffin granule membranes. We now report that calmodulin is bound with high affinity to granule membrane proteins of molecular weights (M_s) 25,000 and 22,000 (25K and 22K) at low Ca^{2+} (<10⁻⁸ M) and to proteins with M_s 69K and 50K at high Ca^{2+} (>1 μM). Other cytosolic components (M_s 70K, 36K, 34K and 32K) require calmodulin for their interaction with membrane. These proteins separately bound to calmodulin-Sepharose at high Ca^{2+} concentrations. Although the functions of these adrenal proteins have not been established, the 34K and 32K M_s components co-migrate with clathrin light chains isolated from medullary coated vesicles and the M_s 34K components from both sources share the same one-dimensional peptide map. These interactions were observed at micromolar Ca^{2+} levels at 'intracellular' conditions of pH and ionic strength and would be expected to occur during secretion from the chromaffin cell.

The drug trifluoperazine has a high affinity for Ca^{2+} -calmodulin and has been shown to inhibit Ca^{2+} -evoked catecholamine secretion from adrenal chromaffin cells^{1,2}. Trifluoperazine also modifies the usual ultrastructural changes in cholinergically stimulated cells and produces an accumulation of unfused chromaffin granules under the plasma membrane³. Direct participation of calmodulin in secretion has been inferred from the blockade of sea urchin cortical granule exocytosis *in vitro* by anti-calmodulin antiserum⁸. The phosphorylation of chromaffin granule membrane proteins is stimulated by Ca^{2+} -calmodulin⁹. These results focus attention on calmodulin as a general transducer of the Ca^{2+} signal in exocytosis-endocytosis, although the biochemistry of its action is unknown.

Calmodulin labelled by ¹²⁵I-Bolton and Hunter reagent (which yields fully biologically active product¹⁰) was used in studies of calmodulin binding to chromaffin granule membranes. High-affinity binding was found even at low (<10⁻⁸ M) Ca^{2+} concentrations. The binding constant ($K_D = 31$ nM) at low Ca^{2+} was the same as previously reported for calmodulin binding to membrane in the presence of Ca^{2+} (10⁻⁴ M)^{4,5}. However, calmodulin membrane binding was sensitive to Ca^{2+} as is clear from the very different values of B_{max} at high and low Ca^{2+} (3.3 and 24 pmol per mg protein), respectively.

Calmodulin-binding proteins in granule membranes were visualized by overlaying¹¹ SDS-polyacrylamide gel electrophoresis (PAGE) gels with the radioactive protein. In the absence of Ca^{2+} , calmodulin was bound exclusively by 22K and 25K M_s components (Fig. 1a). Calmodulin labelled in oxidizing conditions, which impair Ca^{2+} -dependent binding^{10,12}, produced an almost identical labelling pattern⁴. At 1 mM Ca^{2+} , other components in the gel bound calmodulin, principally two

Received 8 September; accepted 16 November 1982.

1. Mudge, A. W., Leeman, S. E. & Fischbach, G. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 526-530 (1979).
2. Konishi, S., Tsunoo, A. & Otsuka, M. *Nature* **282**, 515-516 (1979).
3. Jessell, T. M. & Iversen, L. L. *Nature* **268**, 549-551 (1977).
4. Lightman, S. L., Iversen, L. L. & Forsling, M. L. *J. Neurosci.* **2**, 78-81 (1982).

polypeptides with M_r 69K and 50K (Fig. 1c). Addition of 100 μ M fluphenazine to the calcium solutions prevented this binding (Fig. 1b). In each case the presence of Ca^{2+} decreased calmodulin binding to the 22K and 25K M_r components. Calmodulin-binding sites are apparently oriented towards the cytosol, because when the granule membranes were resealed in conditions permitting active catecholamine uptake¹³, treatment with protease K abolished specific calmodulin binding.

These results, taken together, suggest that at low levels of Ca^{2+} , calmodulin is bound with high affinity to a class of cytosolically oriented granule membrane sites which preferentially bind the Ca^{2+} -free form of calmodulin. In the presence of calcium, calmodulin is bound at other sites and the former interaction becomes quantitatively less important; the apparent affinity constant is not affected.

When chromaffin granule membranes were briefly incubated in adrenal cytosol (100,000g_{av} supernatant of medulla homogenate) several minor cytosolic proteins (M_r 70K, 36K, 34K and 32K) bound reversibly in a calcium-dependent manner^{6,7}. After radiolabelling, the calcium dependence of binding of this group of proteins was found to be closely similar to that of calmodulin itself⁶, with $K_D = 3.6 \mu\text{M}$ Ca^{2+} at 1 mM Mg^{2+} . This interaction was prevented by protease treatment and inhibited by increased concentrations of Mg^{2+} (5–10 mM)⁶.

To determine whether calmodulin itself mediated the interaction, medullary cytosol was depleted of endogenous calmodulin by passage through fluphenazine-Sepharose¹⁴. When calmodulin-depleted cytosol was incubated with membrane, the recovery of the 36K, 34K and 32K components was reduced compared with the untreated control (Fig. 2). Addition of excess calmodulin (10 μ M) to cytosol also reduced binding of the 30K M_r proteins. It cannot be excluded that these proteins were removed together with calmodulin by the fluphenazine column¹⁵. However, specific binding of radiolabelled 32K and 34K proteins to granule membranes in the presence of 100 μ M

Ca^{2+} was low in the absence of calmodulin, increased with up to ~50 nM calmodulin, then decreased again as the calmodulin concentration was raised to 1 μ M (data not shown). Binding could only be abolished completely by addition of EGTA, which may indicate that calmodulin was incompletely removed from the original membranes.

To establish whether the 36K–32K proteins bound directly to calmodulin, cytosolic proteins eluted from granule membrane after incubation with whole cytosol at 10 μ M Ca^{2+} were chromatographed on calmodulin-Sepharose. Proteins with M_r s 70K, 36K, 34K and 32K and trace amounts of other components were eluted with EGTA (Fig. 3). The same column was also used to obtain microgramme amounts of the 34K and 32K proteins, partially purified as described previously⁶, directly from cytosol.

The functions of the novel granule-binding proteins are unknown, although we noted previously⁶ that proteins of M_r s 30K to 36K are present in brain and medulla coated vesicles, where they are tightly associated with clathrin¹⁶. Puszkun and co-workers have recently demonstrated that the two brain clathrin-associated protein 'light chains' (M_r s 33K and 36K) each bind calmodulin after release from clathrin¹⁷. Linden¹⁸ reported that a 32K protein of intact brain coated vesicles was labelled by a photoreactive calmodulin derivative. We purified clathrin coat protein from medulla (Fig. 4) and found that the light chains have a lower M_r than those of brain and co-migrate on SDS-PAGE with the 34K and 32K calmodulin-binding components present in adrenal cytosol (Fig. 5). One-dimensional peptide maps of the 34K M_r medulla clathrin light chain and the corresponding protein from medullary cytosol were the same; lack of material has so far prevented unambiguous interpretation of the 32K protein map (Fig. 6). Despite the apparent identity of the 34K proteins, there must be modifications or structural differences unseen at the present level of resolution, because although adrenal clathrin light

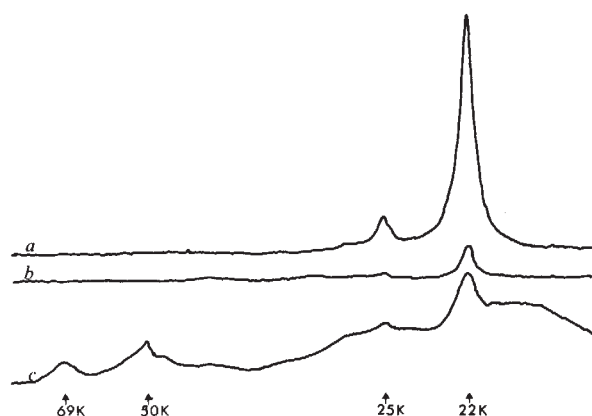


Fig. 1 Densitometry of ^{125}I -calmodulin-binding components of chromaffin granule membrane separated by SDS-PAGE. After electrophoresis of membrane (75 μ g per lane), segments of gel were exposed to 25 μCi ^{125}I -calmodulin in the presence of: a, 1 mM EGTA; b, 1 mM Ca^{2+} + 0.1 mM fluphenazine HCl; c, 1 mM Ca^{2+} . The gel was processed as described by Carlin *et al.*¹¹. Autoradiographs produced via an intensifying screen were densitometered using a soft-laser scanner. As the processing for autoradiography involves extensive washing of the gel¹¹, it is likely that only high-affinity sites are visualized. Bovine chromaffin granules were isolated as described previously⁴, purified by sedimentation through 1.8M sucrose and ultrastructurally homogeneous membrane was prepared by hypo- and hypertonic washes. Membrane was finally resuspended (75 μ g ml^{-1}) in 0.14 M potassium glutamate, 20 mM Na 2[N-morpholino]ethane sulphate (MES), 1 mM Mg^{2+} and 1 mM EGTA, pH 6.5. Calmodulin was iodinated with Bolton and Hunter reagent to a specific activity of 15 μCi per μ g (ref. 10) and diluted 10-fold with unlabelled calmodulin.

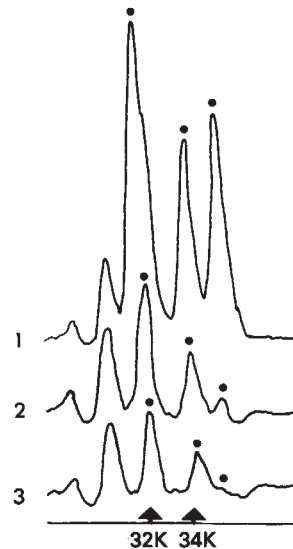


Fig. 2 Ca^{2+} -dependent binding of 30K components of medullary cytosol to chromaffin granule membrane. 1, Initial cytosol; 2, cytosol after passage through fluphenazine-Sepharose; 3, cytosol as in 2 after addition of excess (10 μ M) calmodulin. ● Signifies a calmodulin-binding component. Ca^{2+} -dependent chromaffin granule binding proteins were prepared by incubation of fresh bovine medullary cytosol (20 mg ml^{-1} protein, dialysed against 0.14 M potassium glutamate, 20 mM Na MES, 1 mM Mg^{2+} , 5 mM EGTA- Ca^{2+} and 0.1 mM phenylmethyl-sulphonyl fluoride, pH 5.0, pH 6.5) with chromaffin granule membrane (1 mg ml^{-1}) for 2 min at 37 °C. Adsorbed proteins were separated from cytosol by centrifuging the membrane through a 10% sucrose gradient in the calcium buffer and eluted by resuspending the membrane in the same buffer containing no calcium and re-pelleting the membrane through a calcium-free sucrose gradient. Densitometer traces are from 10% SDS-PAGE separations of eluted proteins.

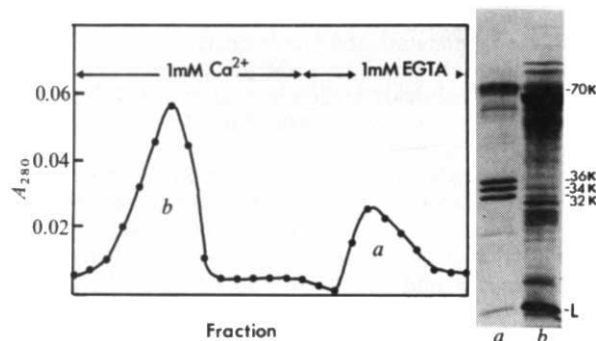


Fig. 3 Calmodulin-Sepharose affinity chromatography of granule-binding proteins prepared as described in Fig. 2 legend. Proteins, dialysed against 1 mM Mg^{2+} , 1 mM Ca^{2+} , 20 mM Tris-HCl pH 8.0, 100 mM NaCl were chromatographed on calmodulin-Sepharose and calmodulin-binding proteins were eluted with 1 mM EGTA. Reduction of the initial Ca^{2+} concentration to 50 μ M also gave the same result. *a*, Bound peak; *b*, unadsorbed peak. L, lysozyme added as a carrier.

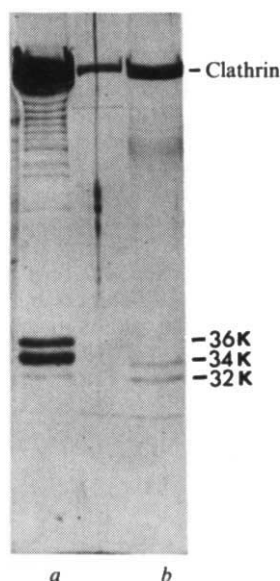


Fig. 4 Clathrin dissociated from brain (*a*) and adrenal medulla (*b*) coated vesicles with 0.5 M Tris. Adrenal clathrin was prepared essentially as described by Keen *et al.*³² for the brain protein.

chains bound to light-chain-free brain clathrin, the cytosolic proteins did not (M. J. G. and E. Ungewickell, unpublished experiments). The structural relationships of these two medullary proteins are being investigated.

The present report describes the nature of both Ca^{2+} -dependent and -independent calmodulin-binding sites on the chromaffin granule membrane. Calmodulin binding may be a general feature of secretory vesicles, as calmodulin is known to bind to synaptic vesicles¹⁹ and has recently been shown to interact with platelet α -granules²⁰. Although calcium-dependent binding properties of calmodulin have received most attention, it is clear that the modulator possesses more than one type of binding site. It is permanently associated, as the δ subunit²¹, with phosphorylase kinase. It is also bound, independently of Ca^{2+} , to a 26K component of the chick eye lens membrane gap junction²². There is a strong analogy with the homologous protein troponin C which complexes both TnT and TnI²³, the latter interaction clearly being Ca^{2+} -dependent. It may be that the muscle system will provide clues to the regulatory roles of calmodulin in non-muscle cells.

The association of calmodulin with accessory proteins normally leads to some process becoming sensitive to regulation

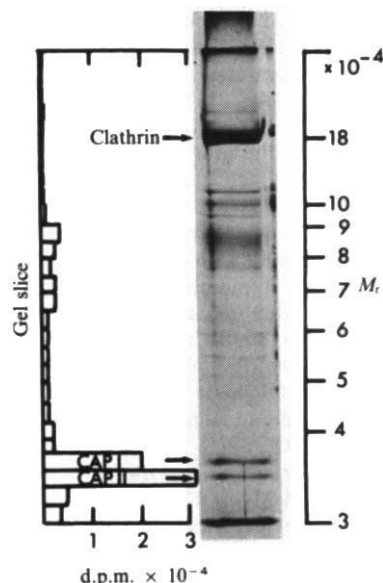


Fig. 5 Co-electrophoresis of 34K and 32K chromaffin granule-binding proteins with medullary clathrin-associated proteins (CAP I and II). 0.1 μ g of ^{125}I -labelled granule-binding proteins was separated on 7% SDS-PAGE. Gel slices (as shown) were counted to localize the 34K and 32K components.

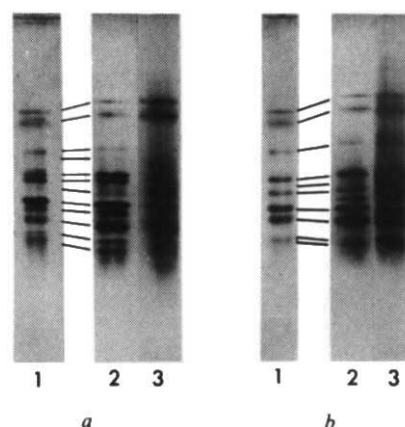


Fig. 6 Limited peptide maps of 32K and 34K proteins obtained by the method of Cleveland³³. The sources of 32K protein (*a*) and 34K protein (*b*) were: 1, medullary cytosol; 2, elution from chromaffin granule membrane after adsorption at 10 μ M Ca^{2+} ; 3, purified clathrin coat protein. Lanes 1 and 3 in each case were silver-stained, lane 2 was stained with Coomassie brilliant blue. The slight difference in relative intensities between lanes 1 and 2 may be the result of the different staining procedure used.

by calcium. Previous reports provide examples of control of enzyme activity (for example, kinases²⁴) or self-assembling systems (for example, microtubules^{25,26} and F actin²⁷). We describe here a new action of the modulator: the recruitment of soluble cytosolic proteins to a subcellular membrane. The concentration of calmodulin in chromaffin cell cytosol²⁸ is sufficient to saturate binding sites on the secretion granules at calcium concentrations in the resting or stimulated cells. Calmodulin-dependent recruitment of the novel cytosolic proteins described occurs at Ca^{2+} concentrations which mediate exocytosis-endocytosis in chromaffin cells^{1,2}. The evidence obtained so far suggests that the recruited proteins might function in endocytosis. Coated vesicles mediate membrane recovery in adrenal medulla²⁹ and both these and secretion granule membranes, following exposure to Ca^{2+} , are associated with a common subset of proteins: calmodulin, a 50K calmodulin-binding membrane protein^{30,31} and two 30K calmodulin-binding components.

We acknowledge the gift of brain triskelions and assays of adrenal medulla light chain binding to brain clathrin, which were performed by Dr E. Ungewickell, and thank Dr Peter Sheterline for drawing the analogy with troponin C.

Received 25 August; accepted 13 November 1982.

- Baker, P. F. & Knight, D. E. *Phil. Trans. R. Soc. B* **296**, 83–103 (1981).
- Knight, D. E. & Baker, P. F. *J. Membrane Biol.* **68**, 107–120 (1982).
- Burgoyne, R. D., Geisow, M. J. & Barron, J. *Proc. R. Soc. B* **216**, 113–115 (1982).
- Burgoyne, R. D. & Geisow, M. J. *FEBS Lett.* **131**, 127–131 (1981).
- Geisow, M. J., Burgoyne, R. D. & Harris, A. *FEBS Lett.* **143**, 69–72 (1982).
- Geisow, M. J. & Burgoyne, R. D. *J. Neurochem.* **38**, 1735–1740 (1982).
- Geisow, M. J. & Burgoyne, R. D. *7th Meet. int. Soc. Neurochem.*, 352 (1981).
- Steinhardt, R. A. & Alderton, J. M. *Nature* **295**, 154–155 (1982).
- Burgoyne, R. D. & Geisow, M. J. *J. Neurochem.* **39**, 1387–1396 (1982).
- Chafouleas, J. G., Dedman, J. R., Munjal, R. P. & Means, A. R. *J. biol. Chem.* **254**, 10262–10267 (1979).
- Carlin, R. K., Orab, D. J. & Siekevitz, P. *J. Cell Biol.* **89**, 449–455 (1981).
- Walsh, M. & Stevens, F. C. *Biochemistry* **17**, 3924–3928 (1978).
- Phillips, J. H. *Biochem. J.* **144**, 311–318 (1974).
- Kakiuchi, S. *et al. FEBS Lett.* **126**, 203–207 (1981).
- Moore, P. B. & Dedman, J. R. *J. biol. Chem.* **257**, 9663–9667 (1982).
- Pearse, B. M. F. *J. molec. Biol.* **125**, 803–812 (1978).
- Lisanti, M. P. *et al. Eur. J. Biochem.* **125**, 463–470 (1982).
- Linden, C. D. *Biochem. biophys. Res. Commun.* **109**, 186–190 (1982).
- De Lorenzo, R. J., Freedman, S. D., Yohe, W. B. & Maurer, S. C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1838–1842 (1979).
- Grinstein, S. & Furuya, W. *FEBS Lett.* **140**, 49–52 (1982).
- Cohen, P. *et al. FEBS Lett.* **92**, 287–293 (1978).
- Welsh, M. J., Aster, J. C., Ireland, M., Alcalá, J. & Maisel, H. *Science* **216**, 642–643 (1982).
- Perry, S. V. *Biochem. Soc. Trans.* **7**, 593–617 (1982).
- Schulman, H., Huttner, W. B. & Greengard, P. in *Calcium and Cell Function* Vol. 1 (ed. Cheung, W. Y.) 219–255 (Academic, New York, 1980).
- Marcum, J. M., Dedman, J. R., Brinkley, B. R. & Means, A. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3771–3775 (1978).
- Kumagi, H. & Nishida, E. *J. Biochem.* **85**, 1267–1271 (1979).
- Sobue, K., Muramoto, Y., Fujita, M. & Kakiuchi, S. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5652–5655 (1981).
- Kuo, I. C. Y. & Coffee, C. J. *J. biol. Chem.* **251**, 1603–1609 (1976).
- Nagasawa, J. & Douglas, W. W. *Brain Res.* **37**, 141–145 (1972).
- Unanue, E. R., Ungewickell, E. & Branton, D. *Cell* **26**, 439–446 (1981).
- Moskowitz, N., Schook, W., Lisanti, M., Hua, E. & Puzsins, S. *J. Neurochem.* **38**, 1742–1747 (1982).
- Keen, J. H., Willingham, M. C. & Pastan, I. H. *Cell* **16**, 303–312 (1979).
- Cleveland, D. W., Fischer, S. G., Kirschner, M. W. & Laemmli, U. K. *J. biol. Chem.* **252**, 1102–1106 (1977).

Phosphorylation of gastrin-17 by epidermal growth factor-stimulated tyrosine kinase

G. S. Baldwin, J. Knesel & J. M. Monckton

Ludwig Institute for Cancer Research,
Melbourne Tumour Biology Unit,
PO Royal Melbourne Hospital, Parkville, Victoria, Australia 3050

Tyrosine phosphorylation seems to be a key event in the control of cellular growth. Several viral transforming proteins, including the src protein of Rous sarcoma virus¹, the p120 protein of Abelson leukaemia virus² and the middle T antigen of polyoma virus³, are phosphorylated by associated tyrosine kinases^{4–6}. The levels of kinase activity correlate with the transforming efficiency of the virus^{6–8}. The receptors for epidermal growth factor (EGF)^{9,10}, platelet-derived growth factor (PDGF)¹¹ and insulin¹² are also phosphorylated by associated tyrosine kinase activities, which are stimulated by EGF, PDGF and insulin, respectively. The EGF-stimulated kinase^{13,14} and the src protein^{14,15} share similar substrate specificity for tyrosines immediately C-terminal to a sequence of acidic amino acids. Such a sequence is also found adjacent to the phosphotyrosine of middle T antigen^{16,17}, and in the homologous region¹⁸ of the hormone gastrin, adjacent to a tyrosine which is sulphated in approximately half the gastrin isolated from gastric mucosa¹⁹. Reports that gastrin acts as a growth factor for cells of the gastrointestinal tract²⁰ suggested that phosphorylation of this tyrosine might be physiologically more relevant than sulphation. We report here that synthetic human gastrin 17 is phosphorylated by the EGF-stimulated tyrosine kinase of A431 cell membranes. The K_m values of 53–87 and 223–547 μ M obtained in the presence and absence of EGF, respectively, are the lowest reported so far for this enzyme.

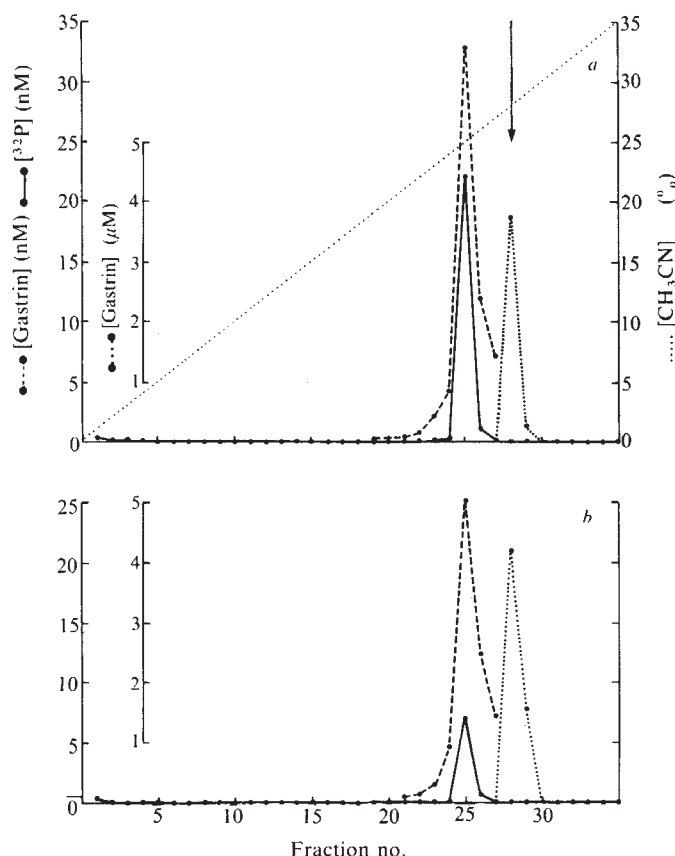


Fig. 1 Reverse-phase HPLC of gastrin after incubation with [γ -³²P]ATP and the EGF-stimulated tyrosine kinase. Synthetic human gastrin-17 (0.17 mM; Research Plus) was preincubated for 5 min at 30 °C in the presence (a) or absence (b) of 2 μ g ml⁻¹ EGF- β (ref. 28) in 50 μ l of 2 mM MnCl₂, 10 μ M Zn(acetate)₂, 0.2% Nonidet-P40, and 20 mM Na⁺-HEPES, pH 7.4 containing 8.4 μ g of A431 cell membranes²¹. Reaction was initiated by adding [γ -³²P]ATP (3.86 Ci mmol⁻¹; 3,827 c.p.m. pmol⁻¹) to a final concentration of 50 μ M, and stopped after 5 min at 30 °C by adding 2.5 μ l of 200 mM EDTA. The reaction mixture was diluted with 1 ml of buffer (0.1 M Na⁺-phosphate pH 6.5, containing 1 mM EDTA) and passed through a Sep-pak cartridge (Waters Associates) which was then washed with 4 $\times$ 5-ml aliquots of buffer A and eluted with 2 ml of 50% CH₃CN in buffer A. The eluate was diluted five-fold with buffer A, and applied to a column (15 cm $\times$ 4.6 mm) of Ultrasphere octadecyl silica (Beckman) which was eluted with the indicated CH₃CN gradient in buffer A at room temperature at a flow rate of 1 ml min⁻¹. The elution position of gastrin-17 in these conditions, determined by A₂₈₀ measurement, is indicated by an arrow. One-minute fractions were collected, and assayed for ³²P (—) by Cerenkov counting and for gastrin-17 (----) by radioimmunoassay³⁰ with an antiserum raised in rabbits against a conjugate of ovalbumin and the gastrin C-terminal tetrapeptide amide. When a purified sample of phosphorylated gastrin (fraction 25, a) was subjected to the above procedure a second time, 73% of the counts applied to the Sep-pak were recovered in the first 2 ml of eluate, and 75% of the counts applied to the HPLC column were recovered in fractions 25 and 26.

Synthetic human gastrin-17 was incubated with a membrane preparation²¹ from the human epidermoid carcinoma cell line A431 (ref. 9) in the presence of EGF, Mn²⁺, Zn²⁺ and [γ -³²P]ATP (Fig. 1). The reaction mixture was adsorbed onto a cartridge of octadecyl silica (Sep-pak; Waters Associates) and gastrin-17 and its phosphorylated derivative were recovered by elution with 50% CH₃CN after most of the ATP and phosphate had been removed by extensive washing of the cartridge. When the Sep-pak eluate was subjected to HPLC on a column of octadecyl silica with a CH₃CN gradient, two peaks of immunoreactive gastrin were observed (Fig. 1). The second

Fig. 2 Ion-exchange HPLC of phosphorylated gastrin. Inset, Identification of phosphotyrosine in acid hydrolysates of phosphorylated gastrin. *a*, 90-h autoradiograph and *b*, markers run on the same plate and located with ninhydrin. PTyr, PThr and PSer represent phosphotyrosine, phosphothreonine and phosphoserine, respectively.

Methods: Phosphorylated gastrin (31,451 c.p.m., 11 pmol) was purified by reverse-phase HPLC (Fig. 1), and applied to a mono-Q column (5 cm × 50 mm; Pharmacia), previously equilibrated with 20 mM Na⁺-phosphate, pH 6.5. The column was washed with 10 ml 0.01% (v/v) acetic acid and eluted with a 50-min gradient from 0.01% to 50% acetic acid at room temperature at a flow rate of 1 ml min⁻¹. The elution position of gastrin-17 in these conditions, determined by A₂₈₀ measurement, is indicated by an arrow. One-minute fractions were collected and assayed for ³²P (—) by Cerenkov counting and for gastrin (----) by radioimmunoassay³⁰ after the acetic acid had been removed by freeze-drying; 36% of the counts applied to the column were recovered in fractions 40–44. Inset, Phosphorylated gastrin (3,534 c.p.m., 1 pmol) was purified by reverse-phase HPLC as described in Fig. 1 legend, and hydrolysed in 6 M HCl at 110 °C for 2 h. After freeze-drying, the hydrolysate was redissolved in 10 µl of H₂O and 5 µl was applied to a 0.1-mm cellulose thin-layer plate (10 × 20 cm; Merck, Darmstadt) before electrophoresis at pH 3.5 as described previously¹⁴.

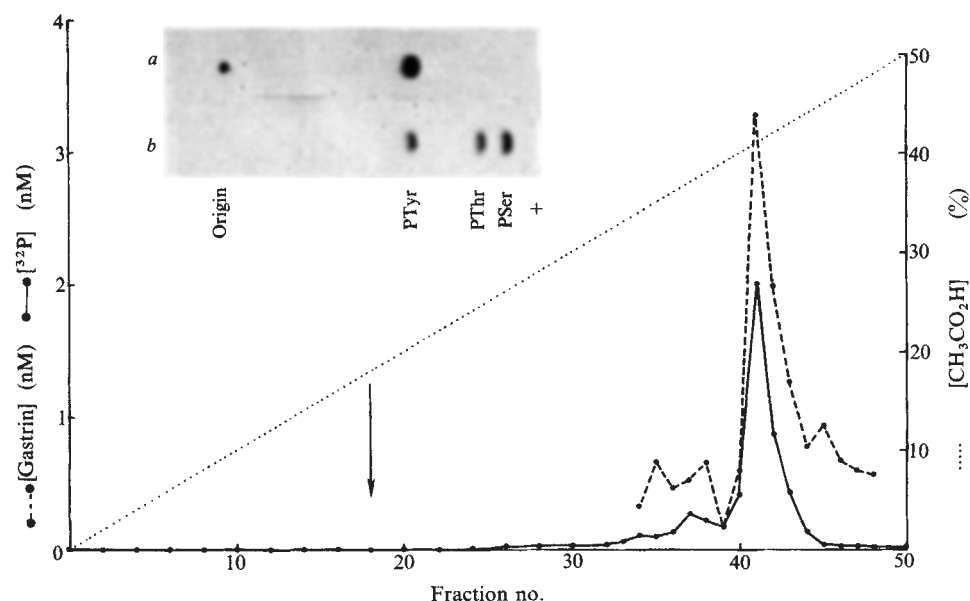


Table 1 Kinetic constants for phosphorylation of gastrin-17

Experiment	Phosphatase inhibitor	$K_m(\pm s.e.)$ (μM)		$V_{max}(\pm s.e.)$ ($pmol\ min^{-1}\ mg^{-1}$)	
		+EGF	-EGF	+EGF	-EGF
1	Zn ²⁺	53 ± 9	223 ± 9	105 ± 7	89 ± 3
2	Zn ²⁺	87 ± 18	547 ± 81	368 ± 34	279 ± 26
3	Zn ²⁺	57 ± 1		169 ± 1	
	VO ₃ ⁻	26 ± 7		206 ± 17	

Gastrin (concentration 43–521 μM) was phosphorylated in the presence and absence of EGF- β (ref. 28) as described in Fig. 1 legend. Phosphorylated gastrin was separated from the other reaction components by adsorption on a Sep-pak cartridge followed by reverse-phase HPLC (Fig. 1), and assayed by Cerenkov counting. All rates were corrected for the 45% loss of phosphorylated gastrin in the purification procedure. K_m and V_{max} values were obtained by fitting the data obtained for three separate preparations of A431 cell membranes to the Michaelis-Menten equation with a weighted least-squares program²⁹. In the third experiment, 0.25 μM NaVO₃ was substituted for 10 μM zinc acetate as a phosphatase inhibitor.

peak eluted at the same position as untreated gastrin-17, whereas the first peak was coincident with a peak of ³²P radioactivity and was therefore assumed to be a phosphorylated form of gastrin-17. This hypothesis was tested by rechromatography of the radioactive material on a mono-Q ion-exchange column with an acetic acid gradient. A single peak of radioactivity was observed, coincident with a peak of gastrin immunoreactivity, and well separated from the elution position of untreated gastrin-17 (Fig. 2). The consistent discrepancy between the concentration of ³²P, calculated from the specific activity of the ATP, and the concentration of gastrin determined by radioimmunoassay, may indicate that the anti-gastrin antibody bound the phosphorylated form of gastrin-17 more tightly than the unmodified molecule. In any case, the coincidence of radioactivity and gastrin immunoreactivity, and the clear separation of the labelled material from unmodified gastrin-17 in two different chromatographic systems, provided compelling evidence for the existence of a phosphorylated form of gastrin-17.

As gastrin-17 contains no histidine, lysine, serine or threonine, the sole tyrosine residue at position 12 seemed the

probable site of phosphorylation. Indeed, acid hydrolysates of phosphorylated gastrin-17 contained phosphotyrosine as the only phosphorylated amino acid (Fig. 2, inset).

EGF stimulated the phosphorylation of gastrin-17 (Fig. 1), by decreasing K_m without significantly affecting the maximum rate of the reaction (Table 1). The K_m for gastrin-17 in the presence of EGF was at least threefold lower than the lowest value reported for phosphorylation of a series of synthetic analogues of residues 414–424 of the src protein by the EGF-stimulated tyrosine kinase¹³. The maximum rate of gastrin-17 phosphorylation was also between 3- and 30-fold lower than the values reported for the same peptides¹³. This difference was not due to inhibition by the Zn²⁺ included to inhibit phosphatase activity²¹, as the same concentration (10 μM) was present in both assay mixtures¹³. Moreover, substitution of NaVO₃ for zinc acetate gave no greater than a twofold change in the kinetic constants for gastrin phosphorylation (Table 1). The observation that the phosphorylation rate for the src peptide in our system²² was similar to published values¹³ suggested that the difference was in fact an inherent property of gastrin-17. The variability in the maximum rates of gastrin-17 phosphorylation in the presence of Zn²⁺ (Table 1) presumably represented variation in the membrane concentration of EGF receptors, due to collection of the cells at slightly different levels of confluence.

The importance of tyrosine modification in the mechanism of gastrin action remains uncertain. Approximately half of the gastrin isolated from gastric mucosa is sulphated on tyrosine¹⁹, but the modification has no apparent effect on binding to receptors on rat gastric mucosal membranes²³, on the serum half-life and dose-response curve for acid secretion²⁴, or on the molecule's ability to stimulate DNA synthesis in gastric mucosa *in vivo*²⁵. However, in the case of analogues of the structurally related hormone, cholecystokinin, the sulphated form of the molecule is ~160-fold more potent than the non-sulphated form in binding to receptors on pancreatic acini²⁶, and in promoting gall bladder contraction²⁷. An indication of the possible significance of gastrin phosphorylation comes from the previously reported sequence homology between gastrin and the middle T antigen of polyoma virus¹⁸. Both molecules may be phosphorylated on a conserved tyrosine. The importance of the phosphorylation of middle T antigen in cellular

transformation by polyoma virus⁶ suggests a possible physiological role for gastrin phosphorylation in the effects of the hormone on growth of gastrointestinal cells.

We thank Mrs Rosemary Chandler for technical assistance, R. H. Whitehead for providing A431 cells, and many other colleagues, particularly A. W. Burgess, B. E. Kemp, E. C. Nice and N. A. Nicola for invaluable discussions.

Received 12 November; accepted 3 December 1982.

1. Czernilofsky, A. P. *et al.* *Nature* **287**, 198–203 (1980).
2. Witte, O. N., Rosenberg, N., Paskind, M., Shields, A. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2488–2492 (1978).
3. Soeda, E., Arrand, J. R., Smolar, N. & Griffin, B. E. *Cell* **17**, 357–370 (1979).
4. Hunter, T. & Sefton, B. M. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1311–1315 (1980).
5. Witte, O. N., Dasgupta, A. & Baltimore, D. *Nature* **283**, 826–831 (1980).
6. Turler, H. *Molec. cell. Biochem.* **32**, 63–93 (1980).
7. Collett, M. S. & Erikson, R. L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2021–2024 (1978).
8. Witte, O. N., Goff, S., Rosenberg, N. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 4993–4997 (1980).
9. Carpenter, G., King, L. & Cohen, S. *J. biol. Chem.* **254**, 4884–4891 (1979).
10. Cohen, S., Carpenter, G. & King, L. *J. biol. Chem.* **255**, 4834–4842 (1980).
11. Nishimura, J., Huang, J. S. & Deuel, T. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4303–4307 (1982).
12. Kasuga, M., Zick, Y., Bliethe, D. L., Crettaz, M. & Kahn, C. R. *Nature* **298**, 667–669 (1982).
13. Pike, L. J., Gallis, B., Casnellie, J. E., Bornstein, P. & Krebs, E. G. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1443–1447 (1982).
14. Hunter, T. *J. biol. Chem.* **257**, 4843–4848 (1982).
15. Smart, J. E. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **78**, 6013–6017 (1981).
16. Neil, J. C., Ghysdael, J., Vogt, P. K. & Smart, J. C. *Nature* **291**, 675–677 (1981).
17. Schaffhausen, B. & Benjamin, T. L. *J. Virol.* **40**, 184–196 (1981).
18. Baldwin, G. S. *FEBS Lett.* **137**, 1–5 (1982).
19. Gregory, R. A. & Tracy, H. J. *Gut* **5**, 103–117 (1964).
20. Johnson, L. R. *Gastroenterology* **70**, 278–288 (1976); **72**, 788–792 (1977).
21. Brautigan, D. L., Bornstein, P. & Gallis, B. *J. biol. Chem.* **256**, 6519–6522 (1981).
22. Baldwin, G. S., Burgess, A. W. & Kemp, B. E. *Biochem. biophys. Res. Commun.* (in the press).
23. Takeuchi, K., Speir, G. R. & Johnson, L. R. *Am. J. Physiol.* **239**, G395–399 (1980).
24. Carter, D. C., Taylor, I. L., Elashoff, J. & Grossman, M. I. *Gut* **20**, 705–708 (1979).
25. Johnson, L. R. & Guthrie, P. D. *Gastroenterology* **71**, 599–602 (1976).
26. Jensen, R. T., Lemp, G. F. & Gardner, J. D. *J. biol. Chem.* **257**, 5554–5559 (1982).
27. Johnson, L. R., Stening, G. F. & Grossman, M. I. *Gastroenterology* **58**, 208–216 (1970).
28. Burgess, A. W., Knesel, J., Sparrow, L. G., Nicola, N. A. & Nice, E. C. *Proc. natn. Acad. Sci. U.S.A.* **79**, 5753–5757 (1982).
29. Kemp, B. E. *J. biol. Chem.* **254**, 2638–2642 (1979).
30. McGuigan, J. E. & Wolfe, M. M. *Clin. Chem.* **28**, 368–373 (1982).

A unique set of polypeptides is induced by γ interferon in addition to those induced in common with α and β interferons

Jon Weil, Charles J. Epstein & Lois B. Epstein

Departments of Pediatrics, Biochemistry and Biophysics and Cancer Research Institute, University of California, San Francisco, California 94143, USA

J. James Sedmak, Jan L. Sabran & Sidney E. Grossberg

Department of Microbiology, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, USA

Human immune interferon (IFN- γ) differs from leukocyte interferon (IFN- α) and fibroblast interferon (IFN- β) in cell origin, inducing agents, physical and biological properties¹ and amino acid sequence². These differences have led to interest in possible differences in the biological properties of IFN- γ compared with IFN- α and IFN- β . IFN- γ has the same broad range of biochemical and biological actions as do IFN- α and IFN- β , although relative potencies vary depending on the cell type and function investigated³. There has so far been no direct evidence that IFN- γ alters normal cell functions differently from other interferons^{1,3}. We report here striking qualitative and quantitative differences in the intracellular response of human fibroblasts to IFN- γ compared with IFN- α and IFN- β . Two-dimensional gel electrophoresis demonstrates, in addition to the induction of a common group of polypeptides, the existence of a set of polypeptides whose synthesis is uniquely induced by IFN- γ .

The effects of IFN- α , IFN- β and IFN- γ on polypeptide synthesis were examined in two human fetal lung fibroblast strains, diploid strain 152 and trisomy-21 strain 153, which have been described previously⁴. Figure 1 shows autoradiograms of two-dimensional polyacrylamide gels from strain 152 cells treated with 200 units ml⁻¹ IFN- γ or with 200 IU ml⁻¹ IFN- α , and from untreated control cultures. (200 IU ml⁻¹ of IFN- α is equivalent to 200 effective antiviral units of IFN- γ .) Each polypeptide whose induction is consistently observed is listed in Table 1. As previously reported, natural IFN- α induces the synthesis of 11 polypeptides that are undetectable in untreated cells^{3,4}. One additional such polypeptide, p51, is now reported here. These polypeptides are also induced by pure IFN- α produced in *Escherichia coli* from cloned DNA (J.W., L.B.E. and C.J.E., unpublished studies). IFN- γ also induces the synthesis of these polypeptides, to a lesser extent (Table 1a) or approximately equal extent (Table 1b) compared with IFN- α .

In addition to the polypeptides strongly induced by IFN- α , IFN- γ induces the synthesis of 12 polypeptides which are little, if at all, affected by IFN- α . These are shown in Table 1c and in Fig. 1 where they are indicated by asterisks. Four of these are undetectable or only marginally detectable in untreated cells, while eight are clearly present in untreated cells. There are thus significant quantitative, and in some instances qualitative, differences in the effects of IFN- α and IFN- γ on polypeptide synthesis in these human fibroblasts. These responses have been observed over a range of IFN- γ concentrations of 2–200 units ml⁻¹.

Based on quantitative estimates of the synthesis of some of the interferon-induced polypeptides, the ratios of the synthesis of polypeptides in cells treated with IFN- γ compared to cells treated with IFN- α are shown in Table 1. Among the polypeptides measured, those which show the greatest differential response to IFN- α and IFN- γ are p44 and the doublet p80b–p81. (Polypeptide p44 is more clearly resolved in some gels than in the gel shown in Fig. 1.) The ratio of induction of the former by IFN- γ compared to IFN- α is ~28. The ratio for the latter is ~0.2. Thus, the synthesis of these polypeptides seems to differ ~140-fold in response to the two types of interferon.

To determine the peptide induction pattern produced by IFN- β , cells of strains 152 and 153 were treated with pure IFN- β (obtained from Dr E. Knight). The pattern of polypeptide synthesis observed was indistinguishable from that obtained with IFN- α .

Several lines of evidence indicate that a contaminant in the IFN- γ preparations is not responsible for induction of the unique set of polypeptides (Table 1c). Each of the effects listed in Table 1 has been observed with four different IFN- γ preparations: crude phytohaemagglutinin-induced and crude concanavalin A-induced (prepared by Dr D. Lucas), partially purified *Staphylococcus enterotoxin A* (SEA)-induced (obtained from Dr M. Wiranowska-Stewart), and highly purified SEA-induced. The last preparation was purified 857-fold by controlled pore glass chromatography followed by lanthanide affinity chromatography. It had a specific activity of 4×10^6 units per mg protein when assayed on A549 cells, a single isoelectric species (pI 8.6), and no detectable IFN- α (J. L. Taylor and J.J.S., unpublished studies). We quantitated five IFN- γ -specific polypeptides (p46, p40, p37, p35 and p28a) and four nonspecific polypeptides (p80a, p78a, p58 and p51) in gels from cells treated with the two crude IFN- γ preparations and with the highly purified preparation. The ratio between the two groups varied less than 50% following induction by the crude preparations and by the highly purified preparation. In addition, cells were treated with a sample of the highly purified preparation that was totally inactivated by dialysis against 0.1M KCl-HCl at pH 2 for 14 days at 4 °C, a treatment which does not inactivate IFN- α or IFN- β . The pattern of polypeptide synthesis following this treatment was indistinguishable from that in untreated cells, thus demonstrating that IFN- α and IFN- β contaminants were not responsible for the induction of

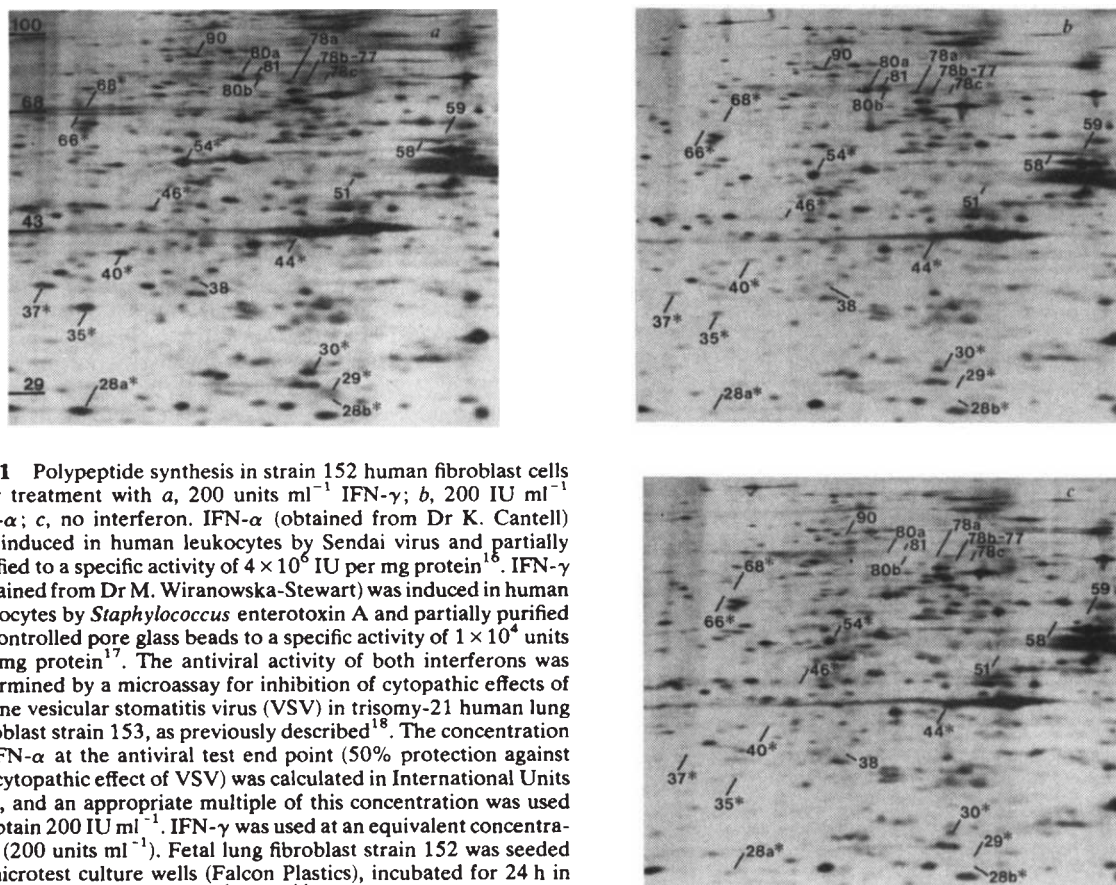


Fig. 1 Polypeptide synthesis in strain 152 human fibroblast cells after treatment with *a*, 200 units ml^{-1} IFN- γ ; *b*, 200 IU ml^{-1} IFN- α ; *c*, no interferon. IFN- α (obtained from Dr K. Cantell) was induced in human leukocytes by Sendai virus and partially purified to a specific activity of 4×10^6 IU per mg protein¹⁶. IFN- γ (obtained from Dr M. Wiranowska-Stewart) was induced in human leukocytes by *Staphylococcus enterotoxin A* and partially purified on controlled pore glass beads to a specific activity of 1×10^4 units per mg protein¹⁷. The antiviral activity of both interferons was determined by a microassay for inhibition of cytopathic effects of bovine vesicular stomatitis virus (VSV) in trisomy-21 human lung fibroblast strain 153, as previously described¹⁸. The concentration of IFN- α at the antiviral test end point (50% protection against the cytopathic effect of VSV) was calculated in International Units (IU), and an appropriate multiple of this concentration was used to obtain 200 IU ml^{-1} . IFN- γ was used at an equivalent concentration (200 units ml^{-1}). Fetal lung fibroblast strain 152 was seeded in microtest culture wells (Falcon Plastics), incubated for 24 h in the presence of 200 $\mu\text{Ci ml}^{-1}$ L-[U- ^{14}C] amino acids (Amersham) with or without interferon, and collected as previously described⁴. Isoelectric focusing followed by 10% SDS-polyacrylamide gel electrophoresis was performed according to O'Farrell¹⁹ and dried gels were exposed on Kodak NS-2T no-screen X-ray film for an equal number of c.p.m.-days⁴. Polypeptides whose concentrations were altered after treatment with one or both interferons are identified by their apparent molecular weights (M_r). Some of these polypeptides were not detectable in untreated cells, and the locations at which they would occur if present are indicated in *c*. Polypeptides that were more strongly induced by IFN- γ than by IFN- α are indicated by an asterisk and are listed in Table 1c. Only that portion of the gels containing interferon-induced polypeptides is shown. The basic side of the isoelectric focusing dimension is to the left. Molecular weight standards were phosphorylase *b*, bovine serum albumin, ovalbumin and carbonic anhydrase, whose locations are indicated on the left-hand side of *a*, plus soybean trypsin inhibitor (21K), whose location is not shown. We have previously shown (ref. 20 and J.W. *et al.*, unpublished data) that the peptide pattern obtained with the partially purified Sendai virus-induced IFN- α used here is identical to that obtained with pure recombinant DNA-produced human IFN- α (IFLrA) and is therefore entirely attributable to IFN- α .

any of the polypeptides. Thus, the induction of the peptides (Table 1a, b), which are strongly induced by both natural and recombinant DNA-produced IFN- α , can be attributed solely to IFN- γ in the IFN- γ preparations and not to other contaminants. Furthermore, in view of the relative constancy of the ratios of the IFN- γ specific to the nonspecific peptides in extracts of cells treated with crude and purified IFN- γ preparations, any putative contaminant responsible for induction of the IFN- γ specific polypeptides would have to quantitatively copurify with IFN- γ several 100-fold through multiple purification steps and would also have to be pH2-labile, as is IFN- γ . We consider the likelihood of this to be extremely small and therefore attribute the induction of the IFN- γ -specific polypeptides directly to IFN- γ .

The gene (*IFRC*) for the species-specific response to interferon has been mapped by interspecific cell hybridization to chromosome 21 (ref. 5), and it has been shown to code for the receptor for IFN- α (ref. 6). As expected from gene dosage considerations, trisomy-21 cells are more responsive to IFN- α than are diploid controls⁴ and are 50% more responsive⁴ to IFN- α -induced polypeptide synthesis. This is consistent with the presence of a third *IFRC* gene copy being responsible for the production of 50% more receptors⁶. However, the sensitivity to highly purified IFN- γ of both the antiviral response and

the induction of polypeptides p80a, p80b-p81 and p58 was greater in the trisomic cells than in the diploid cells (data not shown). These results confirm previous observations made with the partially purified SEA-induced IFN- γ (ref. 3). Since contamination of the highly purified IFN- γ preparation by IFN- α or IFN- β has been ruled out, differential response of trisomy 21 and diploid cells must be attributable to IFN- γ .

Hovanessian *et al.*⁷ used two-dimensional polyacrylamide gel electrophoresis to show that mouse IFN- β and IFN- γ both induce the synthesis of polypeptides having molecular weights of 60,000 (60K) 67K, 88K and 120K. By using one-dimensional polyacrylamide gels, Rubin and Gupta⁸ demonstrated that polypeptides of 56K, 67K, 88K and 120K are induced in human fibroblasts by human IFN- α and IFN- γ but the 67K polypeptide is induced much more strongly by IFN- γ than by IFN- α . In addition, an 80K polypeptide that was induced by IFN- α could not be detected after induction by IFN- γ . The seven polypeptides p81-p77 (Table 1a) we have found may represent a similar result if these polypeptides appeared as a single band in one-dimensional gels.

The results reported here are consistent with IFN- α and IFN- β having a common receptor that differs from that for IFN- γ . Binding competition studies using highly purified ¹²⁵I-labelled recombinant DNA-produced human IFN- α also

Table 1 Characteristics of polypeptides affected by interferon treatment

Designation (molecular weight $\times 10^{-3}$)	Present in untreated cells	Induced by:		Ratio: $\frac{\text{synthesis in cells treated with IFN-}\gamma}{\text{synthesis in cells treated with IFN-}\alpha}$	
		IFN- α	IFN- γ	152 (diploid)	153 (trisomy 21)
<i>a</i> p80a	—	++++	+++	0.15	0.32
p81	—	+++	+++		
p80b	—	+++	+++	0.15	0.27
p78a	—	+++	++	0.26	0.55
p78b	—	++	+		
p77	—	++	+		
p78c	—	++	+		
<i>b</i> p90	—	++	++		
p59	—	++	++		
p58	—	+++	+++	0.57	
p51	—	+++	+++		
p38	—	++	++		
<i>c</i> p68	±	±	+++		
p66	±	±	+++		
p54	+	—	+++	3.8	2.4
p46	+	—	++		2.8
p44	+	—	++++	30.6	26.3
p40	+	±	+++	5.1	5.0
p37	+	+	+++	4.6	4.4
p35	+	±	++++	8.5	6.5
p30	+	±	+		
p29	±	—	+		
p28b	±	—	+		
p28a	+	—	++++		2.8

Polypeptides that are bracketed have identical isoelectric points and similar molecular weights. Polypeptides p78b and p77 are not resolved from one another in Fig. 1. A plus sign (+) indicates that the polypeptide is detectable in gels from untreated cells without any reference to intensity of the spot observed. A minus sign (—) indicates that a given polypeptide is not detectable in untreated cells. ± Indicates that a spot which may represent the polypeptide occurs in high-exposure autoradiograms of gels from some but not all untreated cell preparations. The number of plus signs indicates the relative increase produced by IFN- γ and IFN- α in a given polypeptide as well as the relative effect on different polypeptides, compared with untreated controls. ± Indicates evidence of slight enhancement by IFN- α in some but not all experiments. Cells were treated with 200 IU ml⁻¹ IFN- α (from Cantell) and 200 units ml⁻¹ highly purified SEA-induced IFN- γ (see text and Fig. 1 legend). The intensities of spots on autoradiograms were measured with a scanning microdensitometer and standardized against interferon-invariant control spots as described previously⁴. The polypeptide pair p80b–p81 was scanned as a single unit. Group *a* polypeptides are those induced more by IFN- α than by IFN- γ ; *b* polypeptides are induced to about the same extent by both interferons; *c* polypeptides are induced by IFN- γ and not or only marginally by IFN- α .

indicate that this is the case⁹. A comparable set of interferon receptors has been described in the mouse^{10–13}. All the available data suggest that a gene for the IFN- γ receptor, different from that for the IFN- α and IFN- β receptor, is also located on human chromosome 21. It also appears that the IFN- γ receptor, while different from the receptor for IFN- α and - β , can still trigger the same set of responses as the latter, although with some quantitative differences. However, as IFN- γ also induces significant synthesis of at least 12 additional polypeptides, six of which are not induced by IFN- α or IFN- β , it seems likely that the binding of IFN- γ to its receptor results in the triggering of an additional set of responses.

Equivalent interferon concentrations were used in these experiments based on an antiviral test, using strain 153 cells which are sensitive to all three interferon types. Whereas International Reference Preparations for IFN- α and IFN- β were used to calibrate the assay, a standard for IFN- γ is not yet available (see Fig. 1 legend). The qualitatively different responses of individual polypeptides to IFN- α and IFN- β compared to IFN- γ indicate that the polypeptide results do not simply reflect different interferon concentrations.

The existence of IFN- γ -specific effects on polypeptide synthesis provides a biochemical basis for understanding the reported synergistic interaction between IFN- γ and the other two interferon types¹⁴ and the quantitatively different responses to IFN- γ and IFN- α (ref. 3), including the greatly different dose-response curves of IFN- γ and IFN- α in certain cell types¹⁵. The IFN- γ -specific effects on polypeptide induction further suggest that IFN- γ may have unique biological functions as yet undiscovered.

This work was supported by NIH grants CA-27903, GM-24309, HD-15583 and NO-1-AI-02658, grants 1-760 and 6-126 from the March of Dimes Birth Defects Foundation, and from the Interferon Foundation. We thank Nancy H. McManus for technical assistance and Mary Lou Sumberg and James Harris for editorial assistance.

Received 20 August; accepted 1 December 1982.

- Epstein, L. B. in *Interferon 1981*, Vol. 3 (ed. Gresser, I.) 13–44 (Academic, New York, 1981).
- Gray, P. W. *et al.* *Nature* **295**, 503–508 (1982).
- Epstein, L. B., Weil, J., Lucas, D. O., Cox, D. R. & Epstein, C. J. in *The Biology of the Interferon System* (eds De Maeyer, E., Galasso, G. & Schellekens, H.) 247–256 (Elsevier, Amsterdam, 1981).
- Weil, J., Epstein, L. B. & Epstein, C. J. *J. Interferon Res.* **1**, 111–124 (1980).
- Tan, Y. H., Tischfield, J. & Ruddle, F. H. *J. exp. Med.* **137**, 317–330 (1973).
- Epstein, C. J. *et al.* *Biochem. biophys. Res. Commun.* **107**, 1060–1066 (1982).
- Hovanessian, A. G. *et al.* *Virology* **104**, 195–204 (1980).
- Rubin, B. Y. & Gupta, S. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5928–5932 (1980).
- Branca, A. A. & Baglioni, C. *Nature* **294**, 768–770 (1981).
- Ankel, H., Krishnamurti, C., Besancon, F., Stefanos, S. & Falcoff, E. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2528–2532 (1980).
- Hovanessian, A. G., La Bonnardiere, C. & Falcoff, E. *J. Interferon Res.* **1**, 125–135 (1980).
- Aguet, M. & Blanchard, B. *Virology* **115**, 249–261 (1981).
- Aguet, M., Belardelli, F., Blanchard, B., Marcucci, F. & Gresser, I. *Virology* **117**, 541–544 (1982).
- Fleischmann, W. *Jr Cancer Res.* **42**, 869–875 (1982).
- Grossberg, S. E., Jameson, P. & Sedmak, J. J. in *Handbook of Experimental Pharmacology* (ed. Came, P. & Carter, W.) (Springer, Berlin, in the press).
- Cantell, K. & Hirvonen, S. *J. gen. Virol.* **39**, 541–543 (1978).
- Wiranowska-Stewart, M. *J. Interferon Res.* **1**, 315–321 (1981).
- Epstein, L. B., McManus, N. H., Hebert, S. J., Woods-Hellman, J. & Oliver, D. G. in *Methods for Studying Mononuclear Phagocytes* (eds Adams, D. O., Edelson, P. J. & Koren, H.) 619–628 (Academic, New York, 1981).
- O'Farrell, P. H. *J. biol. Chem.* **250**, 4007–4021 (1975).
- Epstein, L. B., Weil, J. & Epstein, C. J. *UCLA Symp. Molec. Cell. Biol.* **25**, 193–205 (1982).

Tissue-specific expression of a cloned chick δ -crystallin gene in mouse cells

Hisato Kondoh, Kunio Yasuda & T. S. Okada

Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

Crystallins are a group of soluble proteins specific to vertebrate lenses, and have been used successfully as molecular markers for studying lens differentiation. The synthesis and accumulation of δ -crystallin in the differentiating lens cells of chick, in particular, have been extensively studied (for review see refs 1–3). Recently, we have cloned a continuous stretch of a chick δ -crystallin gene which is specifically expressed in lens cells⁴. We have injected this gene into nuclei of various mouse somatic cells. δ -Crystallin is totally absent from mammalian lenses in which it is replaced by γ -crystallin, thus this xenogeneic injection system should allow us to study tissue-specific gene expression. We report here experiments which demonstrate that (1) the cloned δ -crystallin gene of chick is expressed in the mouse lens epithelial cells as efficiently as in the homologous chick cells; (2) δ -crystallin synthesized in the mouse lens epithelium has the native molecular weight, indicating correct splicing of the gene transcripts; (3) the expression is inefficient in the non-lens tissues examined, suggesting lens-specific expression of the δ -crystallin gene in the xenogeneic environment. To our knowledge, this is the first demonstration that a cloned gene shows preferential expression in homologous cell types of different species.

Attempts have been made in several laboratories to clone genes^{4–6} and cDNA sequences^{4,7,8} for δ -crystallin, and analysis of the cloned DNA segments has revealed the presence of at least two non-allelic genes for δ -crystallin in chick^{3–6,8}. We have cloned the entire stretch of one of the δ -crystallin genes in a bacteriophage vector⁴. This gene, which is ~9 kilobases (kb) long and consists of 14 exons and 13 introns⁴, has been re-cloned in a plasmid designated p δ C-1 (Fig. 1).

We reintroduced the cloned δ -crystallin gene into nuclei of mammalian cells to determine whether there was any lens-specific expression. We prepared primary cultures of various mouse tissues and microinjected^{9–11} them with the δ -crystallin gene in the plasmid p δ C-1. The choice of mouse tissues took advantage of the fact that δ -crystallin is replaced by γ -crystallin in mammals and that the two crystallins are not biochemically or immunochemically similar¹, thereby allowing us to identify a very small amount of δ -crystallin synthesized in mouse cells.

The circular p δ C-1 DNA was injected into nuclei of lens epithelial cells in non-growing cultures at a concentration of 60 copies per 10^{-14} l. It has been shown previously that a nucleus can sustain this injection volume^{9–11}. At appropriate intervals after the injection, the cultures were fixed, treated with anti- δ -crystallin antiserum, and processed by the peroxidase-antiperoxidase (PAP) technique¹²; 24 h after the injection, one-quarter of the injected cell population accumulated a detectable amount of δ -crystallin, although these cells showed a considerable variation in staining intensity. Three days after injection, the number of δ -crystallin-positive cells reached a maximum, that is, 40% of the injected cells. Most of the positive cells were heavily stained (Fig. 2a, b). Thereafter, there was a gradual decrease in the number of δ -crystallin-positive cells, reaching 15% by day 6. A similar efficiency has been reported for cells synthesizing viral antigens after nuclear injection of viral DNAs^{9,13}, and a similar time course of transient expression of various exogenous genes has been observed after transfection (see for example refs 14, 15). Injection of the p δ C-1 DNA into the cytoplasm of the same cells yielded no positive cells.

To determine whether the δ -crystallin gene in the mouse lens cells directed the synthesis of authentic δ -crystallin, bulk proteins of the injected cells were analysed electrophoretically to visualize the protein band having δ -crystallin antigenic specificity. The results (Fig. 3) show that the injected cells synthesized a single antigen species of the same molecular weight (48,000) as δ -crystallin of the chick lens epithelium. Thus, it seems that normal δ -crystallin molecules were synthesized in the mouse lens epithelial cells microinjected with the δ -crystallin gene. This implies that in the mouse cells the chick δ -crystallin gene is transcribed and the primary transcripts are spliced correctly.

Figure 3 also shows the abundance of δ -crystallin synthesized in the mouse lens epithelial cells. As judged by the intensity of the δ -crystallin bands (Fig. 3b, d, e), we estimate that the amount of δ -crystallin in the 300 injected mouse cells roughly corresponds to that contained in a few thousands of the chick lens epithelial cells. Assuming that half of the injected cell population actually synthesized δ -crystallin, and taking into account the number of copies of the injected genes per mouse cell and the fact that there are four δ -crystallin genes in a diploid chick cell, we conclude that the δ -crystallin gene is expressed in the mouse lens cells with an efficiency comparable to that in the chick. A separate but similar experiment using a known quantity of δ -crystallin as reference showed that an injected

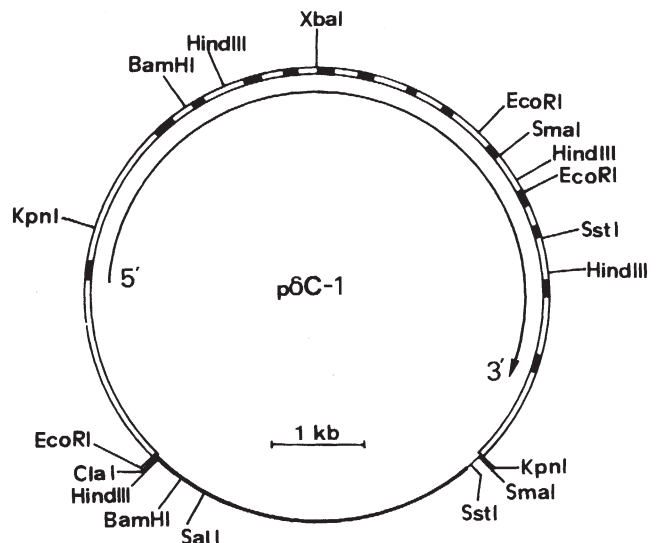


Fig. 1 Structure of p δ C-1. Open and filled boxes indicate noncoding and coding regions of the δ -crystallin gene, respectively. The thin and thick lines represent an SFFV DNA fragment and DNA sequence from plasmid pBR322 (ref. 25), respectively. The direction of transcription of the δ -crystallin gene is indicated by the arrow. Restriction endonuclease sites are also indicated. The DNA segment containing the entire δ -crystallin gene was excised from λ c δ 114 (ref. 4) at the 5'-proximal *EcoRI* site and the 3'-proximal *KpnI* site of the insert, and re-cloned into the plasmid vector pMSF5, replacing the shorter *EcoRI*-*KpnI* segment of the plasmid. pMSF5 is a derivative of pBR322 with an insert at the *PstI* site of the 150-bp DNA segment of SFFV long terminal repeat (LTR) which contains a single *KpnI* site (G. Soma and Y. Ikawa, personal communication). Thus, the plasmid p δ C-1 contains the 3'-proximal 98-bp sequence of the U3 segment including a TATAA box sequence²⁶ and the 5'-proximal 35-bp sequence of the R segment of SFFV LTR located far downstream of the δ -crystallin gene. The LTR segments in the plasmid do not contain any part of the 'activator element' (ref. 27) and the direction of transcription predicted from the orientation of the LTR segment is opposite to that of the δ -crystallin gene. It has been shown that the SFFV LTR of this polarity located at such a position does not activate the associated gene²⁸. It seems highly unlikely that the short LTR fragment contributes much to the expression of the δ -crystallin gene.

Table 1 Expression of the δ -crystallin gene injected into cultured cells of various tissues

Organ	Cell type	Growth	% Injected cells synthesizing δ -crystallin*	No. of injection
Lens	Epithelial	—	40	114
Skin	Fibroblastic	—	2	160
Brain	Epithelial (glial)	—	4	224
Kidney	Fibroblastic	+	0	98
	Epithelial	+	2	200
Lung	Undiscriminated	+	3	211
Liver	Undiscriminated	+	0	301

Mice of strain 129 were used to prepare cultures, which were grown in collagen-coated culture dishes and maintained in Dulbecco's modified Eagle's minimal essential medium containing 10% fetal calf serum. Lens epithelial cells were taken from 10-day-old mice and cultured according to Tsunematsu *et al.*²⁴. DNA was injected 10 days after plating the cells, when the rapid growth phase was over. Cultures of skin fibroblasts and brain glial cells were prepared from an embryo at the 14th day of gestation by trypsinizing abdominal skin and brain, respectively, and were maintained for 4 weeks before receiving DNA. Skin cultures were weekly transferred, whereas brain cultures were kept for the whole culture period without transfer, during which time neuronal cells degenerated and only a sheet of glial cells remained. Cultures of lung, kidney and liver were prepared from a newborn mouse. Injections were done 1 day after inoculating the trypsinized tissues, when several epithelial cells remained. —, Minimal growth; +, rapid growth.

* Assayed 3 days after injection. In cases of growing cultures, each clonal cluster presumably derived from a single injected cells, such as shown in Fig. 1e and f, was counted as one.

mouse lens epithelial cell accumulated on average 0.8 pg or 10^7 molecules of δ -crystallin by 2 days after p δ C-1 injection.

The injection experiment with p δ C-1 was extended to cultures of other mouse tissues (Fig. 2, Table 1). Cells of various non-lens tissues supported the expression of the δ -crystallin gene in only a small fraction (<5%) of the injected cells. The difference in growth potential of cells from different tissues in culture does not seem to be responsible for this phenomenon (Table 1). As judged by staining intensity, the levels of δ -crystallin in some of the positive cells of non-lens tissues seem much lower than those detected in the mouse lens epithelial cells (Fig. 2c, d). It is unlikely that this difference between the lens and non-lens cells is due to different efficiency of DNA injection, as when DNA of pSV2gpt¹⁶ was injected into nuclei of moderately growing cultures of lens epithelium and skin fibroblast, and cultures were labelled with [6-¹⁴C]xanthine for 48 h, ~30% of the injected cells in the two cultures were equally labelled, as judged by autoradiography. Therefore, the greater occurrence of δ -crystallin producers for lens cells must reflect more active expression of the gene, as has been reported for the synthesis of β -globin¹⁴ and Simian virus 40 T-antigen^{13,14,17} after microinjection or transfection of the relevant DNAs. Presumably, the δ -crystallin gene is expressed less efficiently when injected into non-lens tissues, although other factors such as copy number of the plasmid in a nucleus and/or possible chromosomal integration might be responsible also. In this respect, it is interesting that in rapidly growing cultures, as in lung cultures in the present experiments, δ -crystallin-positive cells formed clusters of cells possibly representing clones from injected cells (Fig. 2e, f). As judged by the intensity of their stain, the cells that apparently belonged to a single clone synthesized roughly equal amounts of δ -crystallin, while the amount was variable between the different clones. There is, however, no direct evidence of integration of the injected gene into the chromosome.

In any event, the present results demonstrate that the cloned δ -crystallin gene introduced into mouse cell nuclei is expressed preferentially in the homologous lens epithelial cells. Thus, the lens-specific expression of δ -crystallin genes appears to be mimicked in the xenogeneic mouse cells. This is not, however, due to general enhancement of gene expression in lens epithelial cells. When the whole chick DNA segment 5' of the δ -crystallin

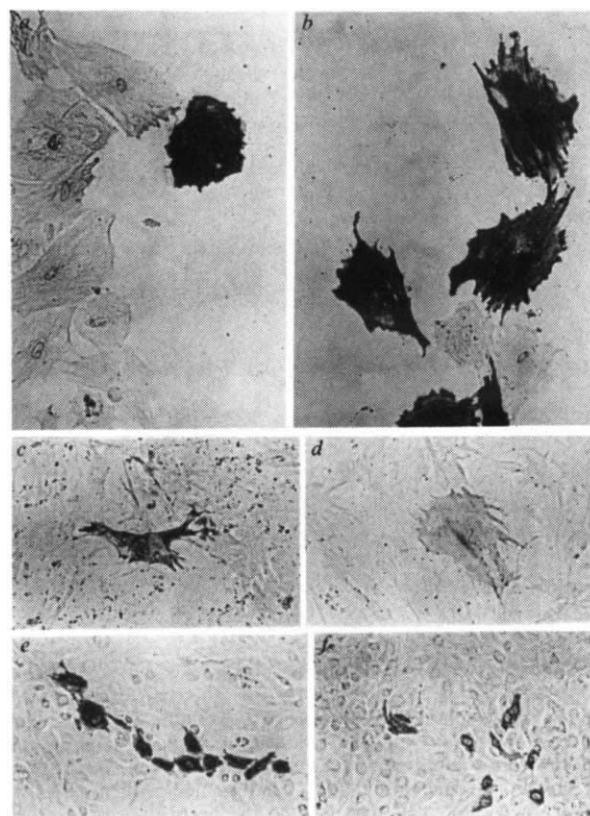


Fig. 2 *In situ* detection of δ -crystallin synthesis in mouse cells after p δ C-1 injection. DNA solution was injected into nuclei using an apparatus described by Yamamoto and Furusawa²⁹ adapted to an inverted microscope. Three days after injection, the cultures were fixed with absolute methanol at -20°C overnight, treated with non-immune swine serum, anti- δ -crystallin rabbit antiserum, anti-rabbit IgG swine serum, and finally horseradish peroxidase-anti-peroxidase rabbit IgG complex (PAP). The peroxidase-catalysed reaction of 1 mM 3-amino-9-ethyl carbazole with 0.01% H_2O_2 was done in 0.1 M sodium acetate pH 5.2 at 37°C for 20 min to stain red the cells containing δ -crystallin-related polypeptides. a, b, Lens epithelial cell: a, a single cell in this area received injection; b, all cells in this area received an injection. c, d, δ -Crystallin-synthesizing cells were only rarely found after p δ C-1 injection into non-growing cultures of skin fibroblasts and brain glial cells, respectively. e, f, Rapidly growing lung fibroblasts. Injected cells formed clones synthesizing abundant (e) and moderate (f) amounts of δ -crystallin. Scale bar, 100 μm .

gene in p δ C-1 was replaced with a putative promoter sequence of Friend spleen focus-forming virus (SFFV) (G. Soma and Y. Ikawa, personal communication), the lens-specific expression of the δ -crystallin gene was not observed, and instead, the gene was expressed nonspecifically in various cultures. For example, with the original gene in p δ C-1, δ -crystallin synthesis in skin fibroblast was less than 5% of the level in lens epithelial cells, when assayed as described in Fig. 3 legend. By contrast, with the recombinant gene plus SFFV promoter, δ -crystallin synthesis was slightly more pronounced in fibroblast culture, that is, 0.2 and 0.4 pg was produced per injection in lens epithelial cells and in skin fibroblasts, respectively. A more detailed account of these experiments will be given elsewhere. Thus, the increased expression in lens epithelial cells is a specific phenomenon dependent on the native promoter region of the crystallin gene. This suggests that the δ -crystallin gene is transcribed faithfully from its promoter in the mouse lens cells.

On the other hand, in view of the fact that the δ -crystallin gene is expressed to some extent in the non-lens cells, its

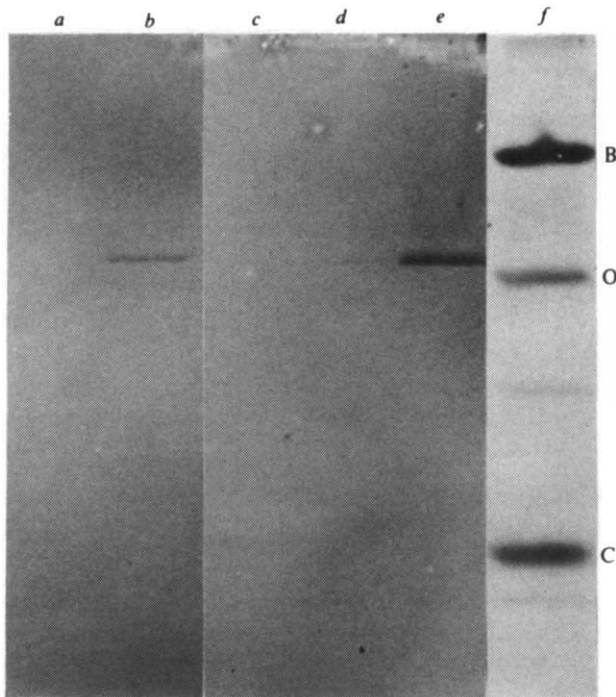


Fig. 3 Electrophoretic detection of δ -crystallin in mouse lens epithelial cells receiving p δ C-1 injection. Bulk proteins in cell cultures were solubilized with SDS, electrophoresed, electroblotted onto a nitrocellulose filter³⁰, and treated with a series of antisera and PAP, and subjected to the peroxidase reaction as described in Fig. 2 legend. *a*, Mouse lens epithelial cell culture containing 10^4 uninjected cells. *b*, A sibling culture containing 300 injected and 10^4 uninjected cells, 2 days after injection. *c*, *d* and *e*, 160, 800 and 4,000 cells, respectively, from a 10-day culture of lens epithelial cells of newly hatched chicks³¹. *f*, Molecular weight markers: B, bovine serum albumin (68,000); O, ovalbumin (45,000); and C, cytochrome *c* (12,400), which were electrophoresed, electroblotted in parallel with *a*–*e*, and stained with amido black.

expression may be under less stringent regulation after transfer to xenogeneic environments compared with that in the native chick cells. This is apparently due to the difference between the state of the δ -crystallin gene in the plasmid on the one hand and in the chromosome on the other. It is possible that some unusual expression of the δ -crystallin gene occurs, albeit with very low efficiency, when the gene in the plasmid is introduced into cell nuclei. For example, transcription initiation could occur upstream of the cloned chick DNA fragment, as was the case with DNA-injected frog oocytes¹⁸ and DNA-transformed mouse cell lines¹⁹, and such transcription might not be regulated normally. It is tempting to assume that this is responsible, at

least in part, for the low level of expression of the δ -crystallin gene after injection of p δ C-1 DNA into non-lens cells.

It seems, however, that normal or nearly normal transcription of the δ -crystallin gene is tissue-specific in the sense that it occurs predominantly in the homologous cell type after injection into the xenogeneic host. Recent studies on viral and specific eukaryotic genes have revealed the presence of DNA regions located upstream from the 5' ends of individual genes that regulate transcription of the respective genes (see, for example, refs 13, 17, 19–23). Our observations with the promoter-replaced δ -crystallin gene suggest the existence of regulatory region(s) 5' of the δ -crystallin gene that are responsible for the lens-specific expression. The DNA injection system reported here should prove useful for identifying regulatory elements on DNA sequences responsible for the tissue-specific expression of genes in cell differentiation.

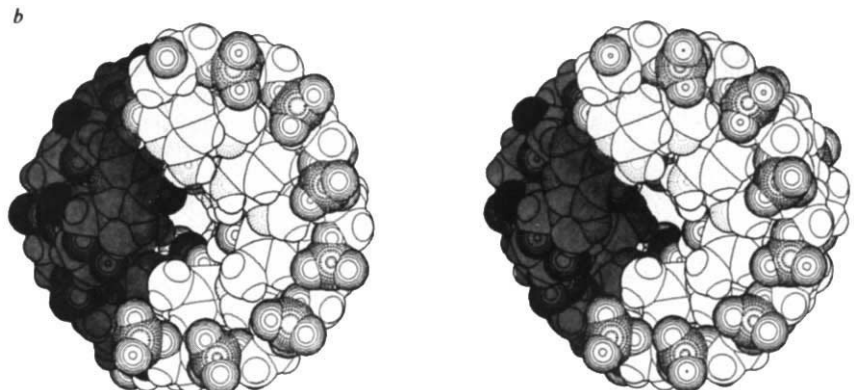
We thank Dr Y. Shimura for helpful discussions and critically reading the manuscript, Drs Y. Ikawa and P. Berg for providing plasmids, S. Hayashi, Y. Shirayoshi and S. Ogoh for assistance, and Ms M. Hase for preparing the manuscript. This work was supported in part by a grant-in-aid for Scientific Research from the Japan Ministry of Education, Science and Culture, and was carried out in accordance with the guidelines of the Ministry for recombinant DNA.

Received 14 September; accepted 25 November 1982.

1. Clayton, R. M. in *The Eye* Vol. 5 (eds Davidson, H. & Graham, L. T.) 399–494 (Academic, London, 1974).
2. Okada, T. S. *Curr. Topics dev. Biol.* **16**, 349–380 (1980).
3. Piatigorsky, J. *Differentiation* (in the press).
4. Yasuda, K., Kondoh, H., Okada, T. S., Nakajima, H. & Shimura, Y. *Nucleic Acids Res.* **10**, 2879–2891 (1982).
5. Bhat, S. P., Jones, R. E., Sullivan, M. A. & Piatigorsky, J. *Nature* **284**, 234–238 (1980).
6. Jones, R. E., Bhat, S. P., Sullivan, M. A. & Piatigorsky, J. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5879–5883 (1980).
7. Bhat, S. P. & Piatigorsky, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3299–3303 (1979).
8. Bower, D. J. *et al. Biochem. J.* **201**, 339–344 (1981).
9. Graessmann, M., Graessman, A. & Mueller, C. *INSERM* **69**, 233–240 (1977).
10. Anderson, W. F., Killos, L., Sanders-Haigh, L., Kretschmer, P. J. & Diakumakos, E. G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5399–5403 (1980).
11. Capecchi, M. *Cell* **22**, 479–488 (1980).
12. Sternberger, L. A. in *Immunocytochemistry*, 104–169 (Wiley, New York, 1979).
13. Moreau, P. *et al. Nucleic Acids Res.* **9**, 6047–6068 (1981).
14. Banerji, J., Rusconi, S. & Schaffner, W. *Cell* **27**, 299–308 (1981).
15. Pellicer, A. *et al. Science* **209**, 1414–1422 (1980).
16. Mulligan, R. C. & Berg, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2072–2076 (1981).
17. Benoist, C. & Chambon, P. *Nature* **290**, 304–310 (1981).
18. Wickens, M. P., Woo, S., O'Malley, B. W. & Gurdon, J. B. *Nature* **285**, 628–634 (1980).
19. Dierks, P., van Ooyen, A., Mantei, N. & Wissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 1411–1415 (1981).
20. Grosschedl, R. & Birnstiel, M. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7102–7106 (1980).
21. McKnight, S. L., Gavis, E. R. & Kingsbury, R. *Cell* **25**, 385–398 (1981).
22. Tsuda, M. & Suzuki, Y. *Cell* **27**, 175–182 (1981).
23. Mellon, P., Gluzman, Y. & Maniatis, T. *Cell* **27**, 279–288 (1981).
24. Tsunematsu, Y., Fukui, H. N. & Kinoshita, J. H. *Exptl Eye Res.* **26**, 671–685 (1978).
25. Boliver, F. *et al. Gene* **2**, 95–113 (1977).
26. Benz, E. W. Jr, Wydro, R. M., Nadal-Ginard, B. & Dina, D. *Nature* **288**, 665–669 (1980).
27. Levinson, B., Khoury, G., Vande Woude, G. & Gruss, P. *Nature* **295**, 568–572 (1982).
28. Joyner, A., Yamamoto, Y. & Bernstein, A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 1573–1577 (1982).
29. Yamamoto, F. & Furusawa, M. *Exptl Cell Res.* **117**, 441–445 (1978).
30. Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350–4354 (1979).
31. Okada, T. S., Eguchi, G. & Takeichi, M. *Dev. Growth Differentiation* **13**, 323–336 (1971).

Erratum

In the article 'Molecular structure of r(GCG)d(TATACGC): a DNA–RNA hybrid helix joined to double helical DNA' by A. H.-J. Wang *et al.*, *Nature* **299**, 601–604 (1982), Figure 2b was poorly reproduced. The correct version is shown here.



Over to you, Mr Andropov

John Erickson

WITH THE recent demise of Leonid Brezhnev and the elevation of Yuri Andropov, erstwhile chairman of the KGB, much speculation has been aroused in East and West about the possibility of our seeing measures sufficiently vigorous and incisive to jerk the Soviet economy out of the stagnation in which it has been long immured. At first sight a KGB background must seem to be most unpromising when it comes to promoting reform of any kind, but even within these somewhat sinister confines Andropov displayed a lively awareness of unsolved problems and the need to deal with them by internal means and resources. Even more pertinently, on returning to the Central Committee Secretariat in May 1982 Andropov almost at once issued a decree demanding an expansion of economic education and wider discussion of current socioeconomic problems in Soviet journals.

Clearly, little in the USSR escapes the attention of the KGB. While it may be going too far to speak of KGB patronage, there is certainly KGB cognizance of the work on management science and computer-based modelling techniques — plus investigation of alternative management systems, price formation mechanisms, flexible wage structures and incentives, all to increase productivity and combat absenteeism — conducted by the research groups in the Academy of Sciences, GOSPLAN and Moscow State University. It goes without saying that science and technology cannot be omitted from the menu of research, not least the intricate problem of exploiting new technology — hence the timeliness and relevance of *Industrial Innovation in the Soviet Union*, even though the book was completed before the advent of Mr Andropov.

The various authors explore a theme which similarly engages Soviet specialists. Here is convenient convergence, for both Soviet and non-Soviet studies tend to agree that the problem is one of severe shortcomings in the organization of both basic and applied research, particularly the arrangement whereby research institutes are subordinated to industrial branch ministries — all designed to create barriers and set up impediments. Consider, for example, proposals set out by V.P. Rassokhin in the journal *Ekonomika i Organizatsiya Promyshlennogo Proizvodstva* (January 1981): Rassokhin argued that certain key industrial research institutes should be detached from their

Industrial Innovation in the Soviet Union. Edited by Ronald Amann and Julian Cooper. Pp.526. ISBN 0-300-02772-9. (Yale University Press: 1982.) £30, \$60.

branch ministries and placed under the State Committee for Science and Technology. The net result would be to establish four systems (hence the title, “*chetvërtaya sistema*”, the fourth system) comprising the Academy of Sciences (for basic research), the higher education system, the existing branch-ministry institutes and a new-found State Committee for Science and Technology “system”, all with the aim of integrating that R&D which went beyond mere branch-industry interest. Predictably Rassokhin’s proposals produced a lively reaction, with general agreement that departmental barriers were a serious problem, but otherwise little sign that a preferred (and workable) solution was in sight.

It is, therefore, neither perversity nor quirkiness which causes me to advise the reader of this massive but excellent work to turn to the last chapter first. This is Julian Cooper’s contribution, “Innovation . . .”, which outlines with a plenitude of footnotes and references the steady Soviet approach to “comprehensive, integrated solutions” for improving R&D, the struggle against “organisational fragmentation”: in effect, reform processes themselves are being systematized and institutionalized, though there is still the problem of the inertia induced by the prevailing organizational structure. While being a general survey, this chapter does delve into specific cases and is all the more valuable for that.

Thus informed and armed, the reader

can turn with profit to the penultimate chapter by Philip Hanson on the role of foreign technology in the Soviet system, a topic which in the present political circumstances is a hot potato. Dr Hanson argues that the overall “weakness in entrepreneurship” within the Soviet system not only hampers domestic innovation but also influences imported technology, not least in impeding the development of an effective “grand strategy for technology import”: in any event, import-led growth was never really a Soviet preference and will decline still further. This chapter must be read and pondered by those who talk glibly of embargoes and the like: in a sense, Dr Hanson is saying that by the very nature of its system the Soviet Union is very often self-embargoed.

To appreciate the wider aims and context of the book it is necessary to turn to Ronald Amann’s penetrating introductory chapter on industrial innovation at large, coupled with a review of the characteristics of the system, the role of history and the influence of the environment. There is indubitably a technology gap with the West, a gap which closes only slowly. Here the central planning system imposes its own limitations on technical progress, but at the same time it is “the main guarantor of defence industry supply priority” — hence the reluctance to interfere with “the system”. In general terms, the defence sector can be accounted something of a success, together with Group Technology, but the price is high. Change, if it is to come, will conceivably depend on “active political support from the highest level” — the emphasis is mine — and the application of the necessary resources, all to breach the

IMAGE
UNAVAILABLE FOR
COPYRIGHT
REASONS

Soviet military technology on show — tanks at Brezhnev’s last parade, November 1982.

resistance (particularly at the lower bureaucratic-administrative levels) to innovation as such. Dr Amann also points to the phenomenon of variations in the technological level of Soviet industry, duly illustrated by the case study approach which occupies Chapters 2-8 and covers the machine-tool industry, Group Technology, the "chemicalisation drive", the management automation programme, control instrumentation for industrial processes and the Soviet defence industry. Here the inclusion of Group Technology and management automation (closely connected with developments in machine tools and computers respectively) represents an addition to the investigations pursued in *The Technological Level of Soviet Industry*, the precursor to the present volume published in 1977.

Dr Amann properly advises on the limitations of the case study approach, though I should echo his point that each of the studies included here is virtually a monograph in its own right. If there is a Soviet success story, then it might be found in John Grayson's elegant and hugely informative study of Soviet Group Technology (GT), the value of which is further enhanced by the tabular presentation of GT outside the USSR, including innovations in Eastern Europe. New also to this volume is the extensive treatment by Dr Cave of the Soviet automated management programme (ASU), which dwells in welcome detail on Soviet computer developments since 1973 and evaluates the problems of the Soviet computer automation programme as a whole. Nevertheless, it might have been useful to have looked more closely at the military automation programme (ASUV) which can fairly be described as a significant aspect of the Soviet "management" scene, if only because *upravlenie* ("management" plus control direction) hides a teeming complexity of issues facing soldier and civilian alike.

The civilian-military interface (or the lack of it) is also germane to David Holloway's two contributions devoted to the defence sector. The first (and more important) of his chapters deals with the system and its structures, especially military R&D organization; the second examines innovation in the field of ICBMs and main battle tanks. Undoubtedly the highest priority goes to military programmes (where innovation can be stimulated by personal interests expressed by the highest leadership level), though the accepted notion of an absolute distinction between military and civilian sectors is open to doubt. This last, a key question, is most admirably and ably explored by Mr Holloway. As for military technology itself, this requires the closest scrutiny and here it would seem that tank technology is not entirely Mr Holloway's forte. Are we talking about design or innovation or both? It should also be noted that there has indeed been a change in the relationship

between the Soviet tank forces and the design/production organization: quite recently Colonel-General (Tank Troops) Yu. Potapov was invested with direct responsibility for R&D on armoured fighting vehicles within the Defence Ministry. This could suggest that all is not well with the T-72 and that there is a demand for a real technology jump with the T-80: moreover, what, if anything is to come of Soviet interest in HIMAG (High Mobility/Agility Vehicle) developments?

This volume is not only impressive but quite invaluable, if only for dispelling myths or combating ignorance. The specialist will no doubt dote on his own section, but the book ought to be used by a wide range of individuals and agencies —

academic, professional or in government (especially the latter). The insights and information are demonstrably significant, the scholarship is formidable; I might however put in a small plea for a composite bibliography, the better to take advantage of key but little known Soviet publications. The further volume which is promised can only be an ornament to this illustrious enterprise, but for the moment the message of the present book is reasonably plain: over to you, Mr Andropov. □

John Erickson is Director of Defence Studies in the Department of Politics, University of Edinburgh. Among other books he is author of Soviet Military Power and Performance (Macmillan, London, 1971).

A New World of ferns and their relations

Peter H. Raven

Ferns and Allied Plants: with Special Reference to Tropical America. By Rolla M. and Alice F. Tryon. Pp.857. ISBN 3-540-90672-X. (Springer-Verlag: 1982.) \$167.80.

THIS magnificently illustrated and beautifully prepared book constitutes a detailed account of the 127 genera of the ferns and other groups of "lower" vascular plants that occur in the Western Hemisphere. It is based on an assessment of the wealth of new information that has become available during the past few decades, as well as on the research of the authors themselves, who have laboured for many years to produce what has turned out to be a masterly synthesis of our knowledge of this group of plants.

The authors adopt a relatively conservative taxonomic treatment with respect to the families of ferns and other vascular plants that do not form seeds, recognizing 33 families (29 of which are treated in this volume), 240 genera and about 9,000 species. They believe that the recognition of more inclusive taxa often reflects relationships better than subdivision to emphasize differences. Since more than half of the genera of ferns are treated in this book in detail, and almost all are discussed to some extent, the work approaches the dimensions of a global survey and will prove invaluable for studies of the group anywhere in the world.

About three-quarters of all ferns occur in the tropics, with nearly half of these (about 3,000 species) in tropical America and an additional 250 species in extra-tropical North and South America. In the Greater Antilles, southern Mexico and Central America, the northern Andean region (with about 1,500 species, or about a sixth of all species of ferns in the world!), and coastal southeastern Brazil, there are striking centres of endemism among the ferns. In contrast, the Amazon Basin is

relatively poor in species of ferns, with only about 300 species in the entire Amazonian region of Brazil, or about a fifth as many as there are in the northern Andean region.

Particularly outstanding is the Tryons' discussion of the spores of ferns, which highlights evolutionary trends, including the line from trilete to monolete apertures and that which leads to the elaborated surface contours derived from the perispores that occur in advanced families.

A treatment of a typical genus in the book includes generic synonymy (except in a few cases where the related Old World genera are especially complex); a description; an overall account of the numbers of species and of the generic subdivisions, if any; a special account and often enumeration of the species found in tropical America; remarks on ecology, geography, spores and cytology; general observations; a distribution map of the genus; supplementary illustrations; and citations literature. The text is supported by handsome black-and-white photographs (taken primarily by Walter H. Hodge) and by scanning electron micrographs of the spores of the various genera.

All of these features add up to make the book a notable contribution to our understanding of this large and often ecologically important group of plants. It certainly deserves a place in every library that attempts to have a collection of the most basic botanical works; in addition many individuals will wish to purchase it, whether or not they have a specialist interest in ferns. One can but applaud and congratulate the authors on having brought this major and original contribution to knowledge to such a satisfactory conclusion. □

Peter H. Raven is Director of the Missouri Botanical Garden and Engelmann Professor of Botany at Washington University, St Louis.

Mostly before Einstein: explanations for the anomalous orbit

P.J. Message

Mercury's Perihelion: From Le Verrier to Einstein. By N.T. Roseveare. Pp.208. ISBN 0-19-858174-2. (Oxford University Press: 1982.) £20, \$49.

INTRODUCTORY accounts of Einstein's general theory of relativity usually place great stress on the importance, amongst observational tests of the theory, of the precession of the orbit of the planet Mercury. Certainly it is a great triumph of that theory that it supplies an explanation of the discrepancy between the observed mean rate of precession of the apse of Mercury and that calculated according to Newton's gravitational theory.

The discrepancy, however, had been known for more than half a century before Einstein's theory was formulated. This book (which is a revision of the author's PhD thesis) follows through the history of the attempts to explain the anomaly, with a thorough survey of relevant publications and discussion of a number of interesting associated topics.

The explanations which were offered fall into two types: that the discrepancy represents the perturbation of Mercury by hitherto undiscovered matter in the Solar System (usually postulated to lie within the orbit of Mercury); or that it indicates that Newton's law is not a completely accurate description of how gravitation actually acts. The general theory of relativity falls into this second type. The further discrepancy (found by Newcomb) between the observed and predicted rates of precession of the orbit plane of Venus at one time also appeared to be significant, and many of the explanations were formulated so as to account for this too. (General relativity of course does not. It is perhaps strange that debate over this second discrepancy virtually ceased between the time when the apse of Mercury had been adequately dealt with by Einstein's theory, and Duncombe's re-analysis of the observations of Venus, published in 1958, which did not confirm the existence of any discrepancy in its motion.)

The earlier explanations, however, were all of the first type. Since changing the estimates of the masses of the known planets could not remove the discrepancy without producing others elsewhere, undiscovered bodies were postulated. It was soon realized that the hypothetical planet "Vulcan", supposed to move in an orbit within that of Mercury, could not be large enough to have given rise to the necessary perturbation otherwise it would have been certainly seen. Similarly the observed level of brightness of the zodiacal light eventually showed that the matter causing it could not be of sufficient mass to

give an adequate explanation of the apparent anomaly in Mercury's orbit.

In the second category of explanation we have "Hall's hypothesis", investigated over some years by Newcomb. Here, the proposed answer was that in the law of gravitation the exponent of the inverse distance is slightly greater than two. Again, this theory was eventually ruled out since it would give effects in the Moon's motion which are not in fact observed. The class of velocity-dependent laws of gravitation, beginning with Weber's, proved equally unable readily to provide the necessary value of the apse motion.

Readers who are confident at mathematical manipulation will no doubt avoid being confused by a number of elementary slips in some of the mathematical explanations given by Roseveare, such as the omission of the largest term in the right-hand side of equation (1.11), by the description of the quantities k_1 and k_2 , in equation (1.12), as "constants" (which they are far from being), and by the use of the notation " $d\varpi$ " for the angle which is traversed between successive passages through the same apse. Such blemishes, however, need not prevent one's appreciating this otherwise thorough study of the fascinating history with which the book is concerned. □

P.J. Message is Reader in Applied Mathematics and Theoretical Physics at the University of Liverpool.

M and O and S

Andrew Holmes-Siedle

MOS Physics and Technology. By E. Nicollian and J.R. Brews. Pp.924. ISBN 0-471-08500-6. (Wiley: 1982.) £61.95, \$99.10.

THE technology of metal-oxide-silicon (MOS) structures has provided much of the thrust which has recently given the Information Revolution its breathtaking pace. Hitherto, a problem for the research worker dealing with MOS devices was that few full, connected accounts of the techniques of MOS physics were available to him. The research literature is immense — my own, highly selective files on MOS research occupy over six feet of shelving — and the interplay of surface physics, semiconductor physics and electronics is subtle.

Now, at last, Bell Laboratories has come

up with the answer and produced another of its gospels. A few months ago, reviewing a textbook emanating from the same establishment, I noted the unique aura of work done there. Nicollian and Brews have produced a book with that same aura — a compound of authority and apparently effortless ease in dealing with a complex, interdisciplinary subject.

Because their aim was to provide unambiguous statements concerning the three layers — metal, oxide, semiconductor — and the crucial interfacial regions between them, the authors have concentrated attention on the simplest possible MOS structure, the MOS capacitor. At the same time, they do not ignore the implications for or needs of the designer of silicon integrated circuits, small chips which include thousands of MOS structures, interwoven with doped (usually diffused) regions. They do, however, omit mention of other semiconductors such as the increasingly important III-V series, probably because MOS technology is not as well-established for these materials.

The MOS capacitor structure in silicon consists of a dot of metal lying on the surface of a silicon dioxide layer, such a layer usually having been formed by thermal oxidation of the underlying silicon in dry oxygen, steam or some similar ambient. The properties of the oxide are extremely sensitive to the method of growth and it might be said that, in the field of high technology, few interacting systems of solids have been as well characterized as the MOS system. This is because of the strong impact which variations in physical structure and purity have on the final electrical characteristics of the working solid-state circuit.

The main tool for investigating those characteristics is the capacitance-voltage plot. As the authors demonstrate, the C-V curve, which can be plotted quickly and automatically, is pregnant with information. Various electrical stresses or other excitations such as electron beams can strongly affect this plot. In separate chapters, the authors describe how to extract from C-V measurements such details as the nature of the interfacial charge traps, interfacial nonuniformities, the doping and carrier lifetime in the silicon and the important and diverse range of network defects which can populate the glassy oxide layer. The versatility and usefulness of the MOS capacitance and conductance techniques are fully worked out and practical experimental details, such as circuits which have not been widely published before, are given. It is thus not surprising that the book is the size of a small bible; for one community of zealots, its international authority may also approach the biblical. □

Andrew Holmes-Siedle is Consultant in Solid State Physics and Radiation Physics to the Fulmer Research Institute, Stoke Poges, Buckinghamshire.

The SCE enigma

Charles H. Ockey

Sister Chromatid Exchange. Edited by Avery A. Sandberg. Pp.706. ISBN 0-8451-2401-3. (Alan R. Liss: 1982.) £74, \$134. *Sister Chromatid Exchange*. Edited by Sheldon Wolff. Pp.306. ISBN 0-471-05987-0. (Wiley: 1982.) £58.20, \$93.10.

IN 1957 J.H. Taylor first demonstrated the phenomenon of sister-chromatid exchange (SCE). Taylor's studies, however, were limited by the poor resolution of ³H-thymidine autoradiography. It was not until 15 years later, and the advent of detection techniques involving differential staining (or fluorescing) of chromatids following uni- and bi-filarly substitution by 5-bromodeoxyuridine, that further progress was made in SCE research.

Using these new techniques it was soon discovered that many known mutagenic and carcinogenic agents increased SCE yields at concentrations 10–100-fold lower than those required to induce other genetical events. Moderate increases in SCEs did not, however, appear to produce any serious perturbations in the cells. These findings suggested that an SCE test would be of great value in discovering even weak genotoxic agents in the environment.

After eight years of intensive research, much of which is documented in the books under review, it may seem surprising that the biological significance and mechanism of induction of SCEs still remain a mystery. Much of the work during this period has concentrated on the relationships between SCE induction and other genetic changes involving the processes of mutation, carcinogenesis, chromosome breakage, meiotic recombination and DNA repair mechanisms. SCEs, however, do not appear to be definitely related to any one of these processes. They are an "S dependent" event induced at the replicating fork as it passes through, or is delayed at, some as yet undefined perturbation or lesion in the DNA helix, possibly at the junctions of adjacent replicons or replication clusters.

The reader of either of these books may also be astonished at the degree to which the various results in this area conflict. For example, the yield of SCEs in peripheral lymphocytes taken from smokers has been found to be either increased, or to remain unaffected; the co-carcinogen, TPA, was found to be an effective SCE inducer by some groups but not by others; and cells from the cancer-prone Bloom's syndrome, which usually manifest a high spontaneous SCE level when grown in the presence of normal cells, have been shown to increase the SCEs in normal cells, while other evidence indicates that the presence of normal cells reduces the SCE levels in the Bloom's cells.

These differences do not necessarily reflect "bad science"; rather, they are a

result of our lack of basic knowledge of the SCE mechanism and the various external factors which influence SCE induction. Some of these factors are discussed in parallel chapters in the two books, particularly in connection with the baseline variation in SCE frequencies in control peripheral blood cultures — SCE induction is very sensitive to environmental conditions during the culture process and is also dependent on the state of health of the lymphocyte donor.

Two other important factors controlling SCE levels, including those in fibroblast cultures and probably *in vivo*, are the time of cell fixation after 5-bromodeoxyuridine incorporation and the concentrations of 5-bromodeoxyuridine used to demonstrate SCEs. The influence of fixation time is still not clear and, again, the results support different conclusions. The other factor, 5-bromodeoxyuridine concentration (abbreviated to BrdU in Sandberg, but to BrdUrd in Wolff) has been used at levels ranging from 0.3 to 200 g/ml. As discussed in both books, BrdU itself is a mutagenic agent which induces SCEs, and as it is required to detect SCEs the presence or absence of a "spontaneous" level still remains a matter for debate. Since these volumes went to press, however, we have been able to show in both peripheral lymphocytes and Chinese hamster fibroblasts that increasing levels of BrdU incorporated into the DNA seriously reduce the expression of SCE lesions in cells treated with certain mutagenic agents. There are also results quoted in both volumes to indicate that the BrdU in the medium has a greater effect on SCE induction than that incorporated into the DNA. As is suggested by authors in each of the books, some form of standardization is required before the SCE method becomes a more reliable weapon in the armoury of tests for genetic hazards in the environment.

The two volumes cover approximately the same general areas of SCE research, and although Wolff contains a more comprehensive index, there are some topics which have been omitted or which are dealt with only in brief. The current theories accounting for SCE induction mechanisms are described by their respective proponents in Sandberg, and this book also includes a lucid chapter discussing the merits of these and other theories. There is some degree of repetition of material both within and between the two volumes, but this is not altogether unsatisfactory; author bias is reduced and the reader is exposed to a healthy mix of controversy and speculation. As to readership, Wolff, the shorter book, covers the most important facets of SCE research and is especially recommended to general and medical readers. Sandberg embraces a broader area and includes more details of experimental test systems, a wider series of test systems and list of drugs tested for SCE induction. This book would therefore be particularly

appreciated by scientists and industrialists.

Between them the two volumes contribute much to our knowledge of this complex yet fascinating corner of biological research. Their publication at this time is opportune, and it is to be hoped that their appearance will encourage further research into unravelling the SCE enigma. □

Charles H. Ockey is Senior Scientific Officer at the Paterson Laboratories, Withington, Manchester.

Naturalist from the Purple Land

Wynford Vaughan-Thomas

W.H. Hudson: A Biography. By Ruth Tomalin. Pp.314. ISBN 0-571-10599-8. (Faber & Faber/Faber Inc.: 1982.) £13.50, \$19.95.

How does the reputation of W.H. Hudson stand today? I suppose that he is a somewhat shadowy figure even to well-informed readers who are keen naturalists. Things were different in the early days of this century. Hudson had become almost a cult figure, the very symbol of freedom and escape in a land where expanding industry was eroding the last traces of the old rural England. Arnold Bennett, then the arbiter of popular literary taste, declared, "How many men know England — the actual earth and flesh of England — as Mr Hudson knows it?"

But the curious fact remains that the man hailed as the authority on the soul of the English countryside never saw it until he was thirty, was the son of an American mother and a father who was half Irish and had spent the whole of his early life on the wide pampas of South America. It was the quarter-Devonian blood that eventually prevailed.

Hudson's early free-ranging life on the pampas, then unfenced and alive with vast flocks of exotic birds, made him an ornithologist — or as he always preferred to call himself, a field naturalist. He gained a well-deserved reputation as a collector of specimens for the Smithsonian Institute in Washington and the Zoological Society of London. Ruth Tomalin has unearthed a photograph of the young Hudson, which shows him as a strikingly handsome man; no wonder his admirers saw him as a "caged eagle" when he eventually settled in London. He found London tough going, and only saved himself from near destitution by marrying his landlady. He was devastatingly honest about his marriage in later years: "Now I was never in love with my wife or she with me, but we became good friends". Emily Hudson was more than a good friend. She believed in

Hudson's genius, and, in spite of poverty, sustained him loyally through the lean years, to the beginning of the new century when Hudson made his breakthrough.

He was helped by two important factors. The turn of the nineteenth century saw the climax of the bicycle boom. For the first time the sensitive, not too well-heeled townsman could escape into a countryside still free from the motor car. Hudson's books on Downland, the West Country and even on the birds in the London parks struck exactly the right note, with their mixture of meticulous observation, deep knowledge and a touch of mystic communication with Nature. As one enthusiast put it, "He will do for England what Thoreau did for America".

Well, perhaps he did, but in a different way. He was the doughtiest warrior in the great Battle of the Birds. His letter to the *Times* on "Feathered Women" became the favourite pamphlet of the Royal Society for the Protection of Birds. But he made the public aware, not only of the threat to birds, but of the dangers to the countryside as a whole. *A Shepherd's Life*, published in 1910, was hailed as a masterpiece. For the rest of his life, Hudson retained a special place among nature writers and set a fashion in nature books which was widely applauded and imitated. Above all, whether they realize it or not, the half a million bird-watchers in modern Britain owe Hudson a deep debt of gratitude. After all, he is the real patron saint of the RSPB.

Will Hudson bear revival? Ruth Tomalin's well-researched and well-written biography sent me back to turn the

Delayed reactions in protein biosynthesis

H.R.V. Arnstein

Protein Biosynthesis in Eukaryotes. Edited by R. Perez-Bercoff. Pp.501. ISBN 0-306-40893-7. (Plenum: 1982.) \$71.40, £37.49.

IN THESE days of shrinking library budgets it becomes increasingly pertinent to question the usefulness of publishing the proceedings of symposia, advanced courses and similar meetings. Usually highly specialized, they tend to be only of restricted and ephemeral interest, although there are of course notable exceptions such as the Cold Spring Harbor Symposia.

This volume on eukaryotic protein biosynthesis is the record of a NATO Advanced Study Institute held in September 1980 at Maratea, Italy. Since many aspects of the topics covered in the 18 chapters of the book are under active investigation in many laboratories, a two-year delay in publication is simply too long, particularly as no attempt appears to have been made to update at least the literature references by means of addenda.

The book is divided into six sections: "The protein synthesizing machinery of eukaryotes"; "On the importance of being spliced"; "On selecting the right messenger"; "Synthesis and processing of proteins"; "Inhibition of protein synthesis at selected levels"; and "Mechanism of regulation and control". This division is to some extent arbitrary and, moreover, there is a certain variation in the quality and significance of the various contributions; several of them deal with topics that are covered in greater depth in recent reviews elsewhere, but on the other hand very few chapters would be useful to non-specialists.

Despite these criticisms, the book will be useful to research workers in the field as it contains a lot of interesting information. Space does not permit any detailed discussion of individual contributions, but I enjoyed particularly the articles on recognition sites in eukaryotic messenger RNAs (M. Kozak), the function of the cap in mRNA translation (A.J. Shatkin) and the cytoplasmic control of protein synthesis (R.J. Jackson). A major gap in the coverage of eukaryotic protein synthesis is the omission of any discussion of the transfer of proteins into and through membranes. The elegant signal hypothesis proposed by Blobel to account for the vectorial synthesis and release of secretory proteins is surely one of the most important developments (for a recent account see P. Walter and G. Blobel's article in *Nature* 299, 691-698). There are also other, more minor, omissions. For example, in the chapter on mRNP particles there is no mention of Greenberg's recent investigations of mRNP structure by cross-linking the proteins to the RNA by uv irradiation.

Finally, I must take issue with a state-

ment in the foreword that the basic mechanism of protein synthesis was "first unravelled in bacterial systems". Those of us who are old enough to remember and young enough not to have forgotten will recall that Zamecnik's group, working with rat liver preparations, discovered the function of both microsomes (membrane-bound ribosomes) and soluble RNA (transfer RNA). Moreover, the first complete nucleotide sequence of a transfer RNA was that established by R. W. Holley for yeast alanine tRNA. In a book on eukaryotic protein synthesis these earlier and important discoveries deserve to be mentioned. □

H.R.V. Arnstein is Head of the Biochemistry Department at King's College, University of London.

More from the Moon

J.W. Head

Planetary Science: A Lunar Perspective. By Stuart Ross Taylor. Pp.481. ISBN 0-942862-00-7. (Lunar & Planetary Institute, Houston, Texas: 1982.) \$42.95 (US), \$54.95 (elsewhere).

TWO YEARS after Apollo 17, S. Ross Taylor published *Lunar Science: A Post-Apollo View*, a book designed to bring some order to our knowledge of the Moon's evolutionary history resulting from the abundance of information derived from the Apollo and Luna missions. In this book, published eight years later, Taylor has attempted to incorporate the results of the exploration of the other planetary bodies into a broader perspective on the origin and evolution of the Solar System, relying heavily on the lunar data as a basis for this new understanding.

Although he writes with an admitted bias towards geochemistry, the author treats his broad subject matter with fairness and evenness. There are excellent chapters on the exploration of the Solar System, on basic planetary processes such as impact cratering and volcanism, and on the nature of planetary crusts, interiors and the stratigraphic record of the planets. Helpful appendices on sources of data and a glossary make the book a valuable introduction to the field. In addition, reflective summaries of where we have been and where we are going capture the unique spirit of lunar and planetary exploration. A chapter on the origin and evolution of the planets provides an especially well-written summary of a very difficult topic, although the author's preference for the double-planet hypothesis for the origin of the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

W.H. Hudson, the "caged eagle".

pages of *Far Away and Long Ago*. I found it compelling reading. *The Purple Land*, with its picture of Argentina, Brazil and Uruguay in the early days, has a contemporary relevance. I am not so sure about those South American romances. Will modern readers respond to their Rima-type fantasies? Whatever may be other people's judgement, I, for one, am grateful to Ruth Tomalin's excellent book for making me take another look at The Field Naturalist from La Plata. □

Wynford Vaughan-Thomas is a journalist and broadcaster.

Moon is not shared by all.

Taylor points out that the wealth of information about the Moon established that body as a cornerstone in the interpretation of the nature and evolution of the other planets. For example, images and samples illustrate the geological, petrological and geochemical nature of the cratering process, and the fact that it was a planetary process of fundamental importance in the early history of the Solar System. Geophysical data establish the structure of the lunar crust and interior. Petrological and geochemical data from lunar samples show that regional to global melting of the outer several hundred kilometres of the Moon (a magma ocean) accompanied accretion. The chronology compiled from dating lunar samples has been combined with photogeologic observations on volcanic and tectonic features to study lithospheric loading and to document the thickening of the lunar lithosphere with time. As Taylor illustrates so well, it is this understanding of many aspects of the Moon that allows us to draw conclusions about the nature and evolution of other planetary bodies.

In addition to the wealth of well-integrated factual information presented in this book, two points emerge. First, Taylor describes our state of knowledge of planetary bodies (other than the Earth and Moon) as being comparable to that of the Moon prior to Apollo. This is a sobering thought because it illustrates how heavily we rely on lunar analogies to interpret the characteristics and history of the other planets. Experience from Apollo showed us that simple extrapolation from the Earth to the Moon often led us astray. Similarly, we must recognize the inadequacy of our knowledge of the planets and interpret data cautiously and plan new missions carefully.

Secondly, the emphasis on the major advances in our knowledge of the Moon gives the false impression that our exploration and study of that body is largely complete. Nothing could be farther from the truth. Orbital geochemical data and high-resolution images of the surface collected during the Apollo mission cover less than 20 per cent of the surface; gravity data exist for much less than half the Moon and the nature of the deep interior is unknown; and we have no geological samples from the lunar farside or from many of the major mare and highland units on the nearside.

There is much to admire in this commendable book. The title, however, seems to imply a greater level of knowledge than currently exists. A concentrated international programme of planetary exploration over the next several decades, of which lunar exploration is a fundamental part, will help to bring our Solar System into a true perspective. □

J. W. Head is a Professor in the Department of Geological Science at Brown University, Providence, Rhode Island.

Green animals in the eighteenth century

James L. Larson

Nature's Second Kingdom. By François Delaporte. Translated by A. Goldhammer. Pp. 266. ISBN 0-262-04066-2. (MIT Press: 1982.) \$20, £16.

EVERYONE who has read at all widely in the natural history, medicine or physiology of the seventeenth and eighteenth centuries has had to come to terms with analogies drawn between plants and animals. Those terms are usually less than favourable. Scientists and historians alike often describe the use of analogy as a source of error or a cause of stagnation, and they imply that only a method based upon observation and experiment is properly scientific.

François Delaporte has come to conclusions diametrically opposed to these in his study of eighteenth-century conceptions of vegetality. At least a part of his success is due to the ways in which he has limited his study, and it is important to understand these limitations from the beginning. He has excluded any consideration of systematics and concentrated instead upon analyses of plant structures and functions found in the work of eighteenth-century French and English botanists. He has concerned himself not so much with results as with the formulation of problems — why and with what assumptions botanists worked out their tactics and methods. And he has confined himself to the intellectual framework proposed by these scientists for themselves, avoiding as much as possible the judgements of retrospective history.

Delaporte suggests that animal physiology provided effective heuristic models for the study of plants only after mechanistic assumptions at the end of the seventeenth century had reduced the structures of plants and animals to equivalent theoretical status. Since it was already well known that plants carried out functions somewhat like those of animals, an abundance of analogical models was available for the interpretation of what was less well understood in plants. So abundant, in fact, that one recurrent problem was to select models in animals appropriate to some essential function in plants. Delaporte's work clearly establishes that differing analogies drawn from animal physiology are the basis of both the analogical and the experimental traditions in botany.

A brief consideration of the chapter on nutrition helps in understanding his argument on this point. In animals nutrition is associated with well-defined organs. The simplicity of plants, however, is disconcerting, and the problem is to discover the mechanisms involved in absorbing, transporting and elaborating food. One group of scientists of the late seventeenth century — Major, Mariotte and Perrault — associated plant nutrition

with a limited set of parts and mechanisms, and took circulation as their model. Throughout the next century the study of plant nutrition revolved around questions generated by this model — absorption, elaboration, respiration and growth. Another way of looking at nutrition avoided the close identification with particular organs and concentrated instead upon a phenomenon common to both plants and animals, the general function of vegetation. Since animals live and grow by means of absorption, excretion and the motion of fluids, and since plants live and grow, analogy suggests that the mechanisms may be the same. In his studies of fluid statics, for example, Stephen Hales hoped to determine plant structures and elucidate the ways in which they functioned, using absorption and perspiration only as general terms of comparison. He proposed, in fact, a working hypothesis, and used experiments to determine whether or not plant parts played the role he assigned to them. This alternative strategy also continued throughout most of the eighteenth century.

This summary does scant justice to the richness of the sources or the complexity of the argument which Delaporte has constructed around them. But it does show, I think, that he has analysed with equal seriousness statements made by both the proponents of analogy and the adepts of observation and experiment. The following chapters on generation and motion follow a similar course. Delaporte examines the analogies, models and comparisons available and the experiments and manipulations that were possible. In reconstructing each series of problems he shows how these problems were reformulated and modified, and he has looked at ways in which lines of investigation twist and turn, and in some cases come to a halt.

Delaporte's book is very French in its choice of source material and in its schematic development, and there will be readers who will find fault with his rigorously symmetrical treatment and with his calendar of saints. Certainly the world of eighteenth-century science looks very different to a specialist on northern Europe. The chapter on generation, for example, does not analyse two of the great German treatises on the subject — Koelreuter's *Vorläufige Nachricht* and Wolff's *Zwo Abhandlungen* — and is much the poorer for the omission. But I think no historian of science will seriously question the importance of this book: Delaporte has restored something of the actual range and complexity of his subject.

James L. Larson is Professor of Scandinavian and Comparative Literature at the University of California, Berkeley.

Mathematics

- DIEUDONNÉ, J. A Panorama of Pure Mathematics. Pp.289. ISBN 0-1225560-2. (Academic: 1982.) \$29.50.
- GOHBERG, I., LANCASTER, P and RODMAN, L. Matrix Polynomials. Pp.409. ISBN 0-12-287160-X. (Academic: 1982.) \$44.
- JACOBS, H.R. Mathematics: A Human Endeavor, 2nd Edn. Pp.649. ISBN 0-7167-1326-8. (W.H. Freeman: 1982.) £12.60.
- LESTER, F.K. and GAROFALO, J. (eds). Mathematical Problem Solving: Issues in Research. Pp.150. Pbk ISBN 0-89168-049-7. (Franklin Institute Press: 1982.) \$14.50.
- ROGERS, C. and SHADWICK, W.F. Bäcklund Transformations and Their Applications. Mathematics in Science and Engineering, Vol. 161. Pp.334. ISBN 0-12-592850-5. (Academic: 1982.) \$44.

Astronomy

- AVENI, A.F. (ed.). Archaeoastronomy in the New World. Proceedings of an International Conference held September 1981, Oxford. Pp.219. ISBN 0-521-24731-4. (Cambridge University Press: 1982.) £16.
- AVENI, A.F. and URTON, G. Ethnoastronomy and Archaeoastronomy in the American Tropics. Annals of the New York Academy of Sciences, Vol. 385. Pp.363. Hbk ISBN 0-89766-160-5; pbk ISBN 0-89766-161-3. (New York Academy of Sciences: 1982.) \$75.
- HEGGIE, D.C. (ed.). Archaeoastronomy in the Old World. Pp.280. ISBN 0-521-24734-9. (Cambridge University Press: 1982.) £20.
- JORDAN, S. The Sun as a Star. Monograph Series on Nonthermal Phenomena in Stellar Atmospheres. Pp.518. (National Aeronautics and Space Administration: 1981.) Np.
- KOPAL, Z. and RAHE, J. (eds). Binary and Multiple Stars as Tracers of Stellar Evolution. Proceedings of the 69th Colloquium of the International Astronomical Union held August–September, 1981, Bamberg. Astrophysics and Space Science, Vol. 98. Pp.468. ISBN 90-277-1436-3. (Reidel: 1982.) Dfl.155, \$67.50.
- MURDEN, J. Astronomy Handbook. Pp.189. ISBN 0-86272-028-1. (Kingfisher: 1982.) £3.95.
- UNDERHILL, A. and DOAZAN, V. B Stars With and Without Emission Lines. Monograph Series on Nonthermal Phenomena in Stellar Atmospheres. Pp.485. (National Aeronautics and Space Administration: 1982.) Np.

Physics

- BARONE, A. and PATERNO, G. Physics and Applications of the Josephson Effect. Pp.529. ISBN 0-471-01469-9. (Wiley: 1982.) £38.50, \$65.75.
- BATES, Sir D. and BEDERSON, B. (eds). Advances in Atomic and Molecular Physics, Vol. 17. Pp.433. ISBN 0-12-003817-X. (Academic: 1982.) \$58.
- BENEDEK, G. and VALBUSA, U. (eds). Dynamics of Gas-Surface Interaction. Proceedings of the International School of Material Science and Technology held July 1981, Italy. Springer Series in Chemical Physics, Vol. 21. Pp.282. ISBN 3-540-11693-1/0-387-11693-1. (Springer-Verlag: 1982.) DM 72, \$32.
- BOARDMAN, A.D. (ed.). Electromagnetic Surface Modes. Pp.776. ISBN 0-471-10077-3. (Wiley: 1982.) £36.50, \$79.95.
- BORGSTEDT, H.U. (ed.). Material Behaviour and Physical Chemistry in Liquid Metal Systems. Pp.548. ISBN 0-306-40917-8. (Plenum: 1982.) Np.
- BREITENLOHNER, P. and DÜRR, H.P. (eds). Unified Theories of Elementary Particles: Critical Assessment and Prospects. Proceedings of the Heisenberg Symposium held July 1981, München. Lecture Notes in Physics, Vol. 160. Pp.217. ISBN 3-540-11560-9/0-387-11560-9. (Springer-Verlag: 1982.) DM 28, \$12.50.
- BURKHARDT, T.W. and VAN LEEUWEN, J.M.J. (eds). Real-Source Renormalization. Topics in Current Physics, Vol. 30. Pp.214. ISBN 3-540-11459-9/0-387-11459-9. (Springer-Verlag: 1982.) DM 60, \$26.70.
- CASAS-VAZQUEZ, J. and LEBON, G. (eds). Stability of Thermodynamic Systems. Proceedings of the Meeting held September 1981, Spain. Lecture Notes in Physics, Vol. 164. Pp.321. ISBN 3-540-11581-1/0-387-11581-1. (Springer-Verlag: 1982.) DM 38, \$16.90.
- DEHESA, J.S., GOMEZ, J.M.G. and ROS, J. (eds). Interacting Bosons in Nuclei. Proceedings of the 4th Tropical School held September–October 1981, Spain. Lecture Notes in Physics, Vol. 161. Pp.209. ISBN 3-540-11572-2/0-387-11572-2. (Springer-Verlag: 1982.) DM 28, \$12.50.
- FRIES, D.C. and ZEITNITZ, B. (eds). Quarks and Nuclear Forces. Springer Tracts in Modern Physics, Vol. 100. Pp.230. ISBN 3-540-11717-2/0-387-11717-2. (Springer-Verlag: 1982.) DM 78, \$34.70.
- GINZBURG, V.L. and KIRZHITS, D.A. (eds). High-Temperature Superconductivity. Pp.364. ISBN 0-306-10970-0. (Plenum: 1982.) \$55.
- GRANDAL, B. (ed.). Artificial Particle Beams in Space Plasma Studies, Vol. 79. Proceedings of a NATO Advanced Research held April 1981, Norway. ISBN 0-306-40985-2. (Plenum: 1982.) Np.
- IBACH, H. and MILLS, D.L. Electron Energy Loss Spectroscopy and Surface Vibrations. Pp.366. ISBN 0-12-369350-0. (Academic: 1982.) \$49.
- JENSEN, J. and KLEITZ, M. (eds). Solid State Protonic Conductors I: For Fuel Cells and Sensors. Danish–French Workshop held December 1981, Paris. Pp.352. ISBN 87-7492-364-1. (Odense University Press, Denmark: 1982.) Dan.Kr.160, \$19.50.
- KENKRE, V.M. and REINEKER, P. Exciton Dynamics in Molecular Crystals and Aggregates. Springer Tracts in Modern Physics, Vol. 94. Pp.226. ISBN 3-540-11318-5/0-387-11318-5. (Springer-Verlag: 1982.) DM 70, \$31.10.

- LAWLEY, K.P. (ed.). Dynamics of the Excited State. Advances in Chemical Physics, Pp.667. ISBN 0-471-10059-5. (Wiley: 1982.) £35.90, \$77.50.
- LLOSA, J. (ed.). Relativistic Action at a Distance: Classical and Quantum Aspects. Proceedings of the Workshop held June 1981, Spain. Lecture Notes in Physics, Vol. 162. Pp.263. ISBN 3-540-11573-0/0-387-11573-0. (Springer-Verlag: 1982.) DM 33, \$14.70.
- MARLOW, W.H. (ed.). Aerosol Microphysics II: Chemical Physics of Microparticles. Topics in Current Physics, Vol. 29. Pp.189. ISBN 3-540-22400-9/0-387-11400-9. (Springer-Verlag: 1982.) DM 65, \$28.90.
- RODI, W. (ed.). Turbulent Buoyant Jets and Plumes. The Science and Applications of Heat and Mass Transfer, Vol. 6. Pp.184. ISBN 0-08-026492-1. (Pergamon: 1982.) £17.50, \$35.
- SITENKO, A.G. Fluctuations and Non-Linear Wave Interactions in Plasmas. International Series in Natural Philosophy, Vol. 107. Pp.262. ISBN 0-08-125051-3. (Pergamon: 1982.) \$25, \$60.
- STARK, H. (ed.). Applications of Optical Fourier Transforms. Pp.545. ISBN 0-12-663220-0. (Academic: 1982.) \$67.50.
- TSIDILKOVSKI, I.M. Band Structure of Semiconductors. International Series in the Science of the Solid State, Vol. 19. Pp.406. ISBN 0-08-021657-9. (Pergamon: 1982.) £45, \$95.
- VANSELOW, R. and HOWE, R. (eds). Chemistry and Physics of Solid Surfaces IV. Springer Series in Chemical Physics, Vol. 20. Pp.496. ISBN 3-540-11397-5/0-387-11397-5. (Springer-Verlag: 1982.) DM 98, \$43.60.
- WOLFENDALE, A.W. (ed.). Progress in Cosmology. Proceedings of the Oxford International Symposium held September 1981, Oxford. Astrophysics and Space Science, Vol. 99. Pp.344. ISBN 90-277-1441-X. (Reidel: 1982.) Dfl.125, \$54.50.

Chemistry

- ABRAMOVITCH, R.A. (ed.). Reactive Intermediates, Vol. 2. Pp.599. ISBN 0-306-40594-6. (Plenum: 1982.) \$59.50.
- BOSCKE, F.L. Inorganic Ring Systems. Topics in Current Chemistry, Vol. 102. Pp.240. ISBN 3-540-11345-2/0-387-11345-2. (Springer-Verlag: 1982.) DM 96, \$42.70.
- CHARALAMBOUS, G. and INGLET, G. (eds). Chemistry of Foods and Beverages: Recent Developments. Pp.348. ISBN 0-12-169080-6. (Academic: 1982.) \$29.50.
- COULSON, C.A. The Shape and Structure of Molecules, 2nd Edn. Oxford Chemistry Series, 30. Pp.94. Hbk ISBN 0-19-855517-2; pbk ISBN 0-19-855518-0. (Oxford University Press: 1982.) Hbk np; pbk £4.95.
- KEESE, R., MÜLLER, R.K. and TOUBE, T.P. Fundamentals of Preparative Organic Chemistry. Pp.149. ISBN 0-85312-396-9/0-85312-450-7/0-470-27522-7. (Ellis Horwood/Halsted Press: 1982.) £15, \$30.15.
- LEHN, J.-M. (ed.). Preparative Organic Photochemistry. Topics in Current Chemistry, Vol. 103. Pp.94. ISBN 3-540-11388-6/0-387-11388-6. (Springer-Verlag: 1982.) DM 48, \$21.40.
- MCCARTHY, G.J., SILBER, H.B. and RHYNE, J.J. (eds). The Rare Earths in Modern Science and Technology, Vol. 3. Proceedings of the 15th Rare Earths Research Conference held June 1981, University of Missouri at Rolla. Pp.558. ISBN 0-306-40919-4. (Plenum: 1982.) \$59.50.
- MCDONALD, D. and HUNT, L.B. A History of Platinum and Its Allied Metals. Pp.450. ISBN 0-905118-83-9. (Europa, London: 1982.) £20.
- SMITH, J.V. Geometrical and Structural Crystallography. Pp.450. ISBN 0-471-86168-5. (Wiley: 1982.) £18.75, \$32.
- VAINSHTEIN, B.K., FRIDKIN, V.M. and INDENBOM, V.L. Modern Crystallography II: Structure of Crystals. Springer Series in Solid-State Sciences, Vol. 21. Pp.433. ISBN 3-540-10517-4/0-387-10517-4. (Springer-Verlag: 1982.) DM 114, \$50.70.

Computer Science

- BALLARD, D.H. and BROWN, C.M. Computer Vision. Pp.523. ISBN 0-13-165316-4. (Prentice-Hall: 1982.) £31.95.
- KREITZBERG, C.B. and SHNEIDERMAN, B. Fortran Programming: A Spiral Approach, 2nd Edn. Pp.435. ISBN 0-15-528015-5. (Harcourt Brace Jovanovich: 1982.) £11.65.
- SCHWEIFEL, H-P. Numerical Optimization of Computer Models. Pp.389. ISBN 0-471-09988-0. (Wiley: 1981.) £12.50, \$34.50.

Earth Sciences

- CONDIE, K.C. Plate Tectonics and Crustal Evolution, 2nd Edn. Pp.310. Hbk ISBN 0-08-028076-5; pbk ISBN 0-08-028075-7. (Pergamon: 1982.) Hbk \$65; pbk np.
- DEACON, G.E.R. and DEACON, M.B. (eds). Modern Concepts of Oceanography. Benchmark Papers in Geology, 61. Pp.385. ISBN 0-87933-390-1. (Hutchinson Ross: 1982.) \$45.
- FLUGEL, E. Microfacies Analysis of Limestones. Pp.633. ISBN 3-540-11269-3/0-387-11269-3. (Springer-Verlag: 1982.) DM 148, \$65.80.
- HOPLEY, D. The Geomorphology of the Great Barrier Reef: Quaternary Development of Coral Reefs. Pp.453. ISBN 0-471-04562-4. (Wiley: 1982.) £45, \$76.50.
- PICKARD, G.L. and EMERY, W.J. Descriptive Physical Oceanography: An Introduction. 4th (SI) Enlarged Edn. Pp.249. Hbk ISBN 0-08-026280-5; pbk ISBN 0-08-026279-1. (Pergamon: 1982.) Hbk £17.50, \$35; pbk £5.95, \$11.95.

SCHEIDEGGER, A.E. Principles of Geodynamics, 3rd Completely Revised Edn. Pp.395. ISBN 3-540-11323-1/0-387-11323-1. (Springer-Verlag: 1982.) DM 165, \$73.30.

SMART, P. and TOVEY, N.K. Electron Microscopy of Soils and Sediments: Techniques. Pp.264. ISBN 0-19-857574-2. (Oxford University Press: 1982.) £45.

St-JULIEN, P. and BELAND, J. (eds). Major Structural Zones and Faults of the Northern Appalachians. Geological Association of Canada Special Paper 24. Pp.280. ISBN 0-919216-20-X. (Geological Association of Canada: 1982.) \$26.50 (members), \$31.50.

TUCKER, M.E. The Field Description of Sedimentary Rocks. Pp.112. Pbk ISBN 0-351-10036-8/0-470-27239-2. (Open University Press/Halsted: 1982.) £4.95.

ZIEGLER, P.A. Geological Atlas of Western and Central Europe. Pp.130 + 40 maps. ISBN 0-444-02084-3. (Elsevier Biomedical/North Holland: 1982.) \$68, Dfl.160.

Biological Sciences

ANTHONY, C. The Biochemistry of Methylotrophs. Pp.385. ISBN 0-12-058820-X. (Academic: 1982.) £24, \$49.50.

BARLOW, H.B. and MOLLON, J.D. (eds). The Senses. Pp.490. Hbk ISBN 0-521-24474-9; pbk ISBN 0-521-28714-6. (Cambridge University Press: 1982.) Hbk £30; pbk £12.50.

BERRIDGE, M.J., TREHERNE, J.E. and WIGGLESWORTH, V.B. (eds). Advances in Insect Physiology, Vol.16. Pp.368. ISBN 0-12-024216-8. (Academic: 1982.) £29.80, \$61.

BODIS-WOLLNER, I. (ed.). Evoked Potentials. Annals of the New York Academy of Sciences, Vol.388. Pp.738. Hbk ISBN 0-89766-166-4; pbk ISBN 0-89766-167-2. (New York Academy of Sciences: 1982.) Hbk np; pbk \$157.

BROWN, R.M. Jr. (ed.). Cellulose and Other Natural Polymer Systems: Biogenesis, Structure, and Degradation. Pp.519. ISBN 0-306-40856-2. (Plenum: 1982.) \$49.50.

CALKINS, J. (ed.). The Role of Solar Ultraviolet Radiation in Marine Ecosystems. Marine Sciences, Vol.7. Pp.724. ISBN 0-306-40909-7. (Plenum: 1982.) \$79.50.

CATSIMPOOLAS, N. (ed.). Cell Analysis, Vol.1. Pp.336. ISBN 0-306-40864-3. (Plenum: 1982.) \$37.50.

DUBNAU, D.A. (ed.). The Molecular Biology of the Bacilli, Vol.1: *Bacillus subtilis*. Pp.378. ISBN 0-12-222701-8. (Academic: 1982.) \$44.

EDWARDS, M.A. and MCDONNELL, U. (eds). Animal Disease in Relation to Animal Conservation. Symposia of the Zoological Society of London, No.50. Proceedings of a Symposium held November 1981, London. Pp.324. ISBN 0-12-613350-6. (Academic: 1982.) £26.40, \$49.

ESSER, K. Cryptogams: Cyanobacteria, Algae, Fungi, Lichens. Pp.610. ISBN 0-521-23621-5. (Cambridge University Press: 1982.) £37.50.

HAMMES, G.G. Enzyme Catalysis and Regulation. Pp.265. ISBN 0-12-321962-0. (Academic: 1982.) \$14.95.

HARRIS, J.R. (ed.). Electron Microscopy of Proteins, Vol.3. Pp.250. ISBN 0-12-327603-9. (Academic: 1982.) £27.80, \$57.

WILSON, L. (ed.). Methods in Cell Biology, Vol.24. The Cytoskeleton, Part A. Cytoskeletal Proteins, Isolation and Characterization. Pp.445. ISBN 0-12-564124-9. (Academic: 1982.) \$49.

Applied Biological Sciences

ALEXANDROV, S.N. Late Radiation Pathology of Mammals. Fortschritte der Onkologie Band 6. Pp.156. (Akademie-Verlag: 1982.) DM 48.

AMERICAN SOCIETY OF PEDIATRIC NEUROSURGERY. Concepts in Pediatric Neurosurgery, II. Pp.222. ISBN 3-8055-3454-X. (Karger: 1982.) SwFr.190, DM 228, \$114.

BLOOM, S.R. and LONG, R.G. (eds). Radioimmunoassay of Gut Regulatory Peptides. Pp.194. ISBN 0-7216-1750-6. (Saunders: 1982.) £10.50.

BONGA, J.M. and DURZAN, D.J. (eds). Tissue Culture in Forestry. Pp.420. ISBN 90-247-2660-3. (Martinus Nijhoff: 1982.) Dfl.115, \$49.50.

BRAIN, J., SCHILL, W.-B. and SCHWARZSTEIN, L. (eds). Treatment of Male Infertility. Pp.330. ISBN 3-540-10990-0/0-387-10990-0. (Springer-Verlag: 1982.) DM 146, \$64.90.

BURGEN, A.S.V. and ROBERTS, G.C.K. (eds). Topics in Molecular Pharmacology, Vol.1. Pp.249. ISBN 0-444-80354-8. (Elsevier/North Holland Biomedical: 1981.) \$74.50, Dfl.160.

CHIARENZA, G.A. and PAPAPOSTOPOULOS, D. (eds). Clinical Application of Cerebral Evoked Potentials in Pediatric Medicine. Proceedings of the International Conference held January 1982, Italy. Pp.415. UK ISBN 90-219-0595-7; US ISBN 0-44-90278-3. (Excerpta Medica/Elsevier/North Holland: 1982.) Dfl.149, \$69.30.

COHN, W.E. (ed.). Progress in Nucleic Acid Research and Molecular Biology, Vol.27. Pp.303. ISBN 0-12-540027-6. (Academic: 1982.) \$37.

CORKIN, S. *et al.* (eds). Alzheimer's Disease: A Report of Progress in Research. Aging, Vol.19. Pp.552. ISBN 0-89004-685-9. (Raven: 1982.) \$59.84.

DOERGE, R.F. (ed.). Wilson and Gisvold's Textbook of Organic Medicinal and Pharmaceutical Chemistry, 8th Edn. Pp.876. ISBN 0-397-52092-1. (Lippincott, Philadelphia: 1982.) Np.

EDWARDS, R.G. and PURDY, J.M. (eds). Human Conception *In Vitro*. Proceedings of the 1st Bourn Hall Meeting. Pp.435. ISBN 0-12-232740-3. (Academic: 1982.) £20, \$49.

ELIAS, J.M. Principles and Techniques in Diagnostic Histopathology: Developments in Immunohistochemistry and Enzyme Histochemistry. Pp.342. ISBN 0-8155-0903-0. (Noyes: 1982.) \$48.

FARRELL, P.M. (ed.). Lung Development: Biological and Clinical Perspectives, Vol.2. Neonatal Respiratory Distress. Pp.307. ISBN 0-12-249702-3. (Academic: 1982.) \$47.50.

FEDTKE, C. Biochemistry and Physiology of Herbicide Action. Pp.202. ISBN 3-540-11231-6/0-387-11231-6. (Springer-Verlag: 1982.) DM 156, \$69.30.

FEFER, A. and GOLDSTEIN, A.L. (eds). The Potential Role of T Cells in Cancer Therapy. Progress in Cancer Research and Therapy, Vol.22. Pp.312. ISBN 0-89004-747-2. (Raven: 1982.) \$50.32.

FOGH, J. and GIOVANELLA, B.C. (eds). The Nude Mouse in Experimental and Clinical Research, Vol.2. Pp.587. ISBN 0-12-261862-9. (Academic: 1982.) \$58.

HATFIELD, J.L. and THOMASON, I.J. (eds). Biometeorology in Integrated Pest Management. Pp.491. ISBN 0-12-332850-0. (Academic: 1982.) \$35.

HILL, D.S. and WALLER, J.M. Pests and Diseases of Tropical Crops, Vol.1: Principles and Methods of Control. Pp.175. ISBN 0-582-60614-4. (Longman: 1982.) £3.30.

HONRUBIA, V. and BRAZIER, (eds). Nystagmus and Vertigo: Clinical Approaches to the Patient with Dizziness. UCLA Forum in Medical Sciences, No.24. Pp.306. ISBN 0-12-355080-7. (Academic: 1982.) \$26.

ISHIZAKA, K. (ed.). Regulation of the IgE Antibody Response. Progress in Allergy, Vol.32. Pp.346. ISBN 3-8055-3470-1. (Karger: 1982.) SwFr.213, DM 256, \$127.75.

JAMIESON, G.A. (ed.). Interaction of Platelets and Tumor Cells. Progress in Clinical and Biological Research, Vol.89. Pp.523. ISBN 0-8451-0089-0. (Alan R. Liss: 1982.) £32.80, DM 144.

Psychology

ACOSTA, F.X., YAMAMOTO, J. and EVANS, L.A. Effective Psychotherapy for Low-Income and Minority Patients. Pp.164. ISBN 0-306-40879-1. (Plenum: 1982.) \$14.95.

GOTTESMAN, I.I. and SHIELD, J. Schizophrenia: The Epigenetic Puzzle. Pp.258. Hbk ISBN 0-521-22573-6; pbk ISBN 0-521-29559-9. (Cambridge University Press: 1982.) Hbk £18; pbk £6.95.

NOTMAN, M.T. and NADELSON, C.C. (eds). The Woman Patient, Vol. 3: Aggression, Adaptations, and Psychotherapy. Pp.314. ISBN 0-306-40859-7. (Plenum: 1982.) \$24.50.

SILVERSTONE, T. and TURNER, P. Drug Treatment in Psychiatry, 3rd Edn. Pp.287. Pbk ISBN 0-7100-9050-1. (Routledge & Kegan Paul: 1982.) £7.95.

TREXLER, L.E. (ed.). Cognitive Rehabilitation: Conceptualization and Intervention. Proceedings of a Conference held January 1981, Indiana. Pp.280. ISBN 0-306-41016-8. (Plenum: 1982.) \$37.50.

History of Science

DOWNS, R.B. Landmarks in Science: Hippocrates to Carson. Pp.305. ISBN 0-87287-295-5. (Libraries Unlimited, Colorado: 1982.) \$23.50 (US), \$28 (elsewhere).

HALL, A.R. From Galileo To Newton. Pp.363. ISBN 0-486-24227-7. (Dover, New York: 1982.) £4.50.

HOLTON, G. and ELKANA, Y. (eds). Albert Einstein: Historical and Cultural Perspectives. The Centennial Symposium held March 1979, Jerusalem. Pp.439. ISBN 0-691-08299-5. (Princeton University Press: 1982.) £24.70.

WARD, P. The Adventures of Charles Darwin. Pp.96. ISBN 0-521-24510-9. (Cambridge University Press: 1982.) £3.95.

Anthropology

HAMMOND, N. Ancient Maya Civilization. Pp.337. Hbk ISBN 0-521-24017-4; pbk ISBN 0-521-28399-X. (Cambridge University Press: 1982.) Hbk £22.50; pbk £7.95.

SABLÖFF, J.A. Archaeology: Myth and Reality. Readings from Scientific American. Pp.101. Hbk ISBN 0-7167-1394-2; pbk ISBN 0-7167-1395-0. (W.H. Freeman: 1982.) Hbk £12.80; pbk £6.30.

SCHIFFER, M.B. (ed.). Advances in Archaeological Method and Theory, Vol.5. Pp.478. ISBN 0-12-003105-1. (Academic: 1982.) \$42.

General

ARYA, V.K. and AGGARWAL, J.K. (eds). Deconvolution of Seismic Data. Benchmark Papers in Electrical Engineering and Computer Science, Vol.24. Pp.308. ISBN 0-87933-406-1. (Hutchinson Ross: 1982.) \$45.

CLAYTON, G.D. and CLAYTON, F.E. (eds). Patty's Industrial Hygiene and Toxicology, 3rd Edn, Vol. 2C Toxicology with Cumulative Index for Vol.2. Pp.5,112. ISBN 0-471-09258-4. (Wiley: 1982.) £75, \$130.

EISENSTEIN, E.L. The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe. Pp.794. Pbk ISBN 0-521-29955-1. (Cambridge University Press: 1982.) £12.50.

EMDE, R.N. and HARMON, R.J. (eds). The Development of Attachment and Affiliative Systems. Pp.311. ISBN 0-306-40849-X. (Plenum: 1982.) \$22.50.

EMRICK, H.M., ALDENHOFF, J.B. and LUX, H.D. (eds). Basic Mechanisms in the Action of Lithium. Proceedings of a Symposium held October 1981, Bavaria, FRG. Pp.265. UK ISBN 90-219-0520-5; US ISBN 0-444-90249-X. (Excerpta Medica/Elsevier/North Holland: 1982.) Dfl.175, \$81.50.

HAND, D.J. Kernel Discriminant Analysis. Pp.253. ISBN 0-471-10211-3. (Wiley: 1982.) £15.50, \$34.95.

LASS, N.J. (ed.). Speech and Language: Advances in Basic Research and Practice, Vol.6. Pp.477. ISBN 0-12-608606-0. (Academic: 1981.) \$55.

LOWRY, R.J. (ed.). The Journals of Abraham Maslow. Abridged Edn. Pp.365. Hbk ISBN 0-86616-015-9; pbk ISBN 0-86616-016-7. (Lewis Publishing Co: 1982.) Hbk np; pbk \$12.95.

MACSWEENEY, D. The Crazy Ape: Sanity, Madness, Your Brain and You. Pp.244. ISBN 0-7206-0565-2. (Peter Owen, London: 1982.) Np.

MEAD, C. and SMITH, K. Hertfordshire Breeding Bird Atlas. Pp.128. ISBN 0-9507951-0-0. (HBBA, Herts: 1982.) £5.